DOCUMENT RESUME

ED 335 383 TM 016 996

TITLE	Standardized Tests and Our Children: A Guide to Testing Reform = Los Examenes Estandarizados y Nuestros Ninos: Guia para la Reforma de Examenes Estandarizados.
INSTITUTION	National Center for Fair and Open Testing (FairTest), Cambridge, MA.
SPONS AGENCY	Edward W. Hazen Foundation, Inc., New Haven, CT.; Lilly Endowment, Inc., Indianapolis, Ind.
PUB DATE	90
NOTE	65p.
AVAILABLE FROM	FairTest, 342 Broadway, Cambridge, MA 02139 ($3.00, 5 for $10.00, 10 for $15.00, inquire for larger orders).
PUB TYPE	Guides - Non-Classroom Use (055) -- Multilingual/Bilingual Materials (171)
LANGUAGE	English; Spanish
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Accountability; Achievement Tests; Criterion Referenced Tests; Educational Assessment; *Educational Change; Elementary Secondary Education; Intelligence Tests; Norm Referenced Tests; Performance; School Readiness Tests; Screening Tests; *Standardized Tests; *Student Evaluation; Student Promotion; *Testing Problems; Test Use
IDENTIFIERS	*Authentic Assessment

ABSTRACT

English and Spanish versions of a guide to testing reform are presented. The guide explains alternatives to standardized testing for educational assessment and shows how these new methods of evaluation will improve education. Standardized tests are used for an increasing number of purposes in the school system, but there is no proof that their use has improved schools. Common uses/types of standardized tests include: (1) readiness tests; (2) developmental screening tests; (3) intelligence tests; (4) tracking; and (5) grade promotion. Norm-referenced and criterion-referenced tests are discussed, and the flaws inherent in standardized testing are reviewed. Improved ways of evaluating students are briefly discussed, including performance assessments. Suggestions are provided to help promote authentic assessment in the schools. An appendix lists resources available to persons interested in testing reform. The Statement of the Campaign for Genuine Accountability in Education is included. The Campaign maintains that heavy emphasis on standardized, multiple-choice tests is undermining progress toward needed educational reforms in the United States. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

Standardized Tests and Our Children:

A Guide to Testing Reform

by

FairTest

The National Center for Fair & Open Testing

BEST COPY AVAILABLE

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

ROBERT A. SCHAEFFER

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM016996

FairTest

National Center for Fair & Open Testing

Winter, 1990

Dear Reader:

Standardized tests play a major role in our children's education. Scores from these exams can determine the schools they will attend, whether they are "tracked" for college prep or vocational courses, and what special classes they must take. Test results may also be used to label children as "gifted," "learning disabled," or "educable mentally retarded."

Widespread use of standardized tests began in the 1960's. Their use grew even more in the 1980's because many people thought they were a good way to measure school quality and how well our children were learning.

More recently, parents, teachers, and education experts have come to see that these tests do more harm than good. Too often, multiple-choice and short-answer standardized tests provide false information about our children. Many are unfair to students who come from minority or low income families or to those who do not speak English as their first language. Some tests are also sex (gender) biased.

What's more, overuse of standardized tests has hurt the quality of education. Teachers are forced to teach the narrow skills these exams measure — and to teach them in a narrow way. Teachers end up ignoring the "higher order thinking skills" children really need to learn.

In fact, testing has become a major roadblock to needed school improvements. The testing explosion has also taken control of our schools away from local residents and turned power over to large testing companies. As a result, parents have less, not more, useful information about their children's education.

Across the country, more than one hundred million standardized tests are given in public schools each year. That means a typical student must take about three dozen before leaving high school. Many children are required to take their first standardized test as early as age three or four.

Now, a growing number of parents, teachers, civil rights leaders, educational administrators, and other advocates have begun working together to change this situation. Groups have formed in many places to stop the misuse and overuse of flawed tests. Classrooms, schools, districts, and states across the nation are making changes in how we evaluate students. New assessment methods can help teachers be better educators and can provide improved information about student progress and the quality of our schools.

The movement to improve assessment — and with it our nation's schools — has only just begun. This guide will help you understand standardized tests and learn what's wrong with them. Just as importantly, it will present ideas about what you can do to change the way we judge our children and our schools.

We look forward to working with you to make schools better for all our children.

The National Center for Fair & Open Testing (FairTest)

342 Broadway, Cambridge, Mass. 02139 (617) 864-4810 FAX (617) 497-2224

TABLE OF CONTENTS

I. **Introduction** 2

II. **How Are Standardized Tests Used?** 3
A Brief History of Testing
Some Standardized Tests

III. **What Are Standardized Tests — How Do They Work?** 8
What Test Scores Mean

IV. **What's Wrong with Standardized Tests?** 11
What Rights Do Parents Have About Testing Their Children?

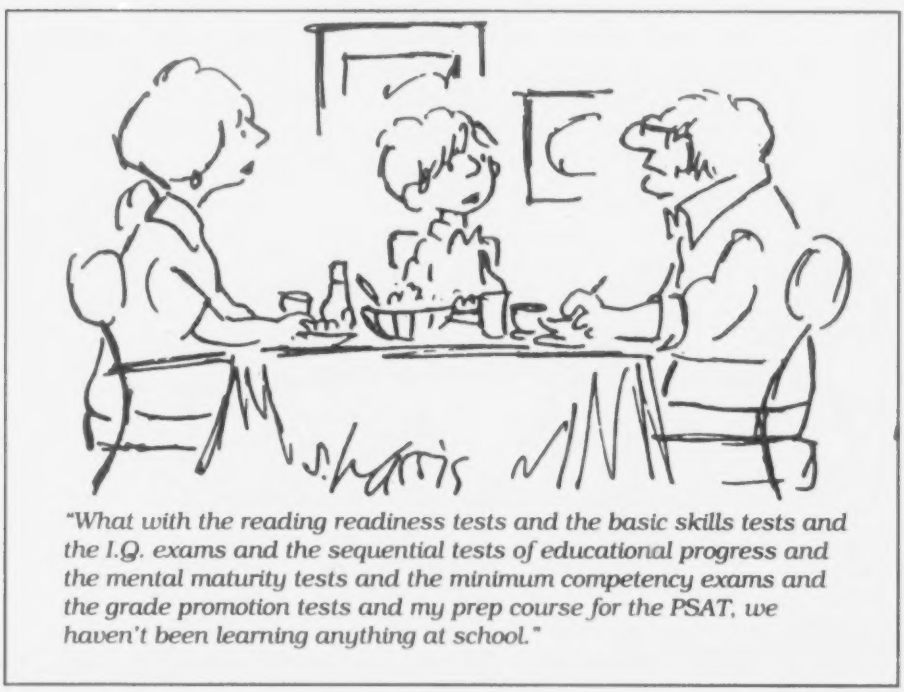
V. **Better Ways to Evaluate Students** 19

VI. **What You Can Do** 22

Appendix: Resources You Can Use 25

Appendix: What Testing Terms Really Mean (Glossary) 26

Statement of FairTest's National Campaign for Genuine Accountability in Education 27



"What with the reading readiness tests and the basic skills tests and the I.Q. exams and the sequential tests of educational progress and the mental maturity tests and the minimum competency exams and the grade promotion tests and my prep course for the PSAT, we haven't been learning anything at school."

BEST COPY AVAILABLE

I. INTRODUCTION

✎ *Mike is barely five years old. This week he took his first standardized test — a test to decide whether he is "ready" for kindergarten. The test said he is "not ready." Now Mike and his parents and teachers all believe he is a failure. But he's really a normal little boy made nervous by a test.*

✎ *Tamika is seven. She just scored below the thirtieth percentile on a multiple-choice, standardized reading test. Her teacher says she is making good progress in school, but her test scores say she should repeat the second grade. If she stays back, she will do the very same work all over again. Tamika says she is bored in school.*

✎ *Kim is ten. He has been held back once, in third grade, and put into a class for "slow learners" because of his test scores. It is April, almost time for the achievement tests. He has been practicing test-taking skills since September. Kim's teacher says he doesn't pay attention. Kim says he hates school because, "All we do is stupid work sheets. School makes me feel dumb."*

✎ *Jimmy is fifteen. Everyone in his community is amazed at his artistic creativity and the unique and interesting ways he views the world. But an IQ test said he was "below normal." So he was tracked into a "slow learner" group, never got good grades, and was held back once. Jimmy plans to drop out of school as soon as he turns sixteen.*

✎ *Juanita is seventeen. She has always earned good grades. This week she received the Scholastic Aptitude Test scores she needs to apply to most colleges. She scored high in math but low in verbal. Because her first language is Spanish, Juanita has had to work extra hard to write well in English, but she took the time in school and it showed. The SAT gave her no chance to think about which word to choose or to carefully correct her work. Her verbal scores were so low she may not get into a good school and has little chance to win a scholarship.*

Testing has hurt all these children. Some have been placed or "tracked" into dead end classes that guarantee they will never get to college. For too many, school has become a boring and frustrating place. Even those who do well in school, like Juanita, may find out that tests hurt their chances for success.

These examples are not unusual. Standardized, multiple-choice and short-answer tests hurt many children. This booklet will show you what a standardized test is, what test scores really mean, and how they are used and misused. It will also explain how and why these tests harm education even when they are used as their makers intended.

Fortunately, there are better ways to find out how our children are doing in school. This booklet explains these new methods of evaluation and shows how they will improve our schools.

Many parents, teachers, school administrators, and concerned members of the community have discovered that too much standardized, multiple-choice testing is destructive. In many places, schools, districts, and even states have begun to stop or reduce multiple-choice testing and implement better ways to evaluate student learning.

We can make schools a place where all children learn and are happy. Changing testing is one necessary part of improving education. Together, we can make it happen.

II. HOW ARE STANDARDIZED TESTS USED?

In recent years, standardized testing has expanded greatly. Much of this growth is a response to justified concerns of parents and others that schools are not doing a good job of educating many of our children. Standardized tests are supposed to tell us how well students are learning. The theory is that test results can be used to make teachers and administrators do a better job.

Though more and more standardized tests are given each year, there's no proof that our schools have improved. In fact, federal programs like the National Assessment of Educational Progress have shown there has been almost no improvement in student learning since 1974, even though there is a lot more testing.

Tests are used for an increasing number of purposes. Some of the more common uses are:

Readiness Tests

These tests are used to decide if a child is "ready" to enter kindergarten or first grade. Children judged "unready" are often kept out of school, made to repeat a year, or put in a "transitional" kindergarten or first grade.

Unfortunately, research shows that scores on readiness tests are wrong as much as half the time. Part of the problem is poor quality tests with bad questions. Children also may develop in spurts and don't always learn things in the same order. A child who is labeled "unready" today could be "ready" next month. Boys and children from low-income or minority families are most likely to be labelled "not ready."

Evidence shows that very few students benefit from being held out of school or kept an extra year in a "transitional program." It's simply not true that large numbers of students are "not ready" for school. The real problem is that schools are forcing children to do the wrong kinds of academic work at too young an age — in part to boost test scores in higher grades. Readiness tests are then used to see which children can meet these unreasonable expectations.

If schools are made ready for all our children — instead of children having to be "ready" for school — then there would be no need for readiness testing.

HI AND LOIS

Reprinted with special permission of King Features Syndicate, Inc.

BEST COPY AVAILABLE

Developmental Screening Tests

School administrators are supposed to use "developmental screening tests" on young children only to find out who should be referred to a specialist to see if they need extra help. Researchers have found that these test scores are often wrong. The tests are frequently misused — without parent or teacher input — to place children in programs, not simply to refer them for more study. Using such tests also encourages teachers and mental health workers to focus only on children's weaknesses and ignore their strengths.

Teachers do need to know how to look for signs that a child may need special help, but there are better methods than using most screening tests.

IQ Tests

Schools also use "intelligence quotient" (IQ), "cognitive abilities" and "aptitude" tests to sort and label children when they begin school and again as they enter high school. IQ tests assume that intelligence is a single thing, rather than a variety of abilities. They also assume intelligence is fixed and permanent. Many researchers disagree.

A Few of the Standardized Tests Used in the U.S.

Hundreds of different standardized tests are administered in the nation's schools each year. Parents are rarely given any explanation of a test's purpose. In many cases, teachers, principals and other education officials don't even know what the scores are supposed to mean. Below are only a few of the tests most commonly used:

Readiness Tests
- Gesell School Readiness Test, Gesell Preschool Test
- Metropolitan Readiness Test
- Brigance (several school readiness tests)

IQ, Ability and Aptitude Tests
- Wechsler Intelligence Scales for Children – Revised (WISC-R)
- Stanford-Binet Intelligence Scale
- Otis-Lennon School Ability Test (OLSAT)
- Cognitive Abilities Test (COGAT)
- Test of Cognitive Skills (TCS)
- Kaufman Assessment Battery for Children (K-ABC)
- Woodcock-Johnson Psycho-Educational Battery
- Differential Aptitude Test (DAT)

Screening Tests
- Developmental Indicators for the Assessment of Learning – Revised (DIAL-R)
- Denver Developmental Screening Test (DDST)
- Early Screening Inventory (ESI)
- McCarthy Screening Test
- Minneapolis Preschool Screening Instrument (MPSI)

K – 12 Achievement Tests
- California Achievement Test (CAT)
- Iowa Test of Basic Skills (ITBS)
- Comprehensive Test of Basic Skills (CTBS)
- Metropolitan Achievement Test (MAT or Met)
- Stanford Achievement Test (SAT; SAT also refers to a college admissions test)

Studies have shown that IQ scores can be changed by test coaching, by eating a good breakfast, or simply by having more friendly people administer the test. This suggests that IQ tests are just one type of achievement test. At least two federal courts have ordered schools to stop using IQ tests to place children in educational programs because the tests are racially and culturally biased.

Since there is no evidence that IQ tests can predict how well someone will do later in life, there is no good reason to require all children to take these tests.

Tracking

One test score can decide a child's entire educational career. School personnel often use tests to place or "track" students into different levels of instruction, sometimes called "ability groups." Depending on scores, students will be assigned to a gifted, accelerated, normal, remedial, or special education program.

Tracking is supposed to help students by placing them in work groups that are not too hard or too easy for them. Educational researchers have concluded, however, that tracking almost always hurts children at the lower levels and usually does not help those placed in the top tracks to do better than they would if placed in mixed-level groups.

In a tracked system, higher track students are more likely to work on projects, read books, and discuss issues. Because their schooling is more interesting and meaningful, those in higher tracks learn more and faster. Schools often have lower expectations of students who score low on standardized exams. These students are likely to receive an inferior education. Raising test scores becomes the main focus of their schooling. As a result, much of the classwork in lower tracks is based on memorizing and repeating the information teachers think students need in order to get better test scores. Lower track students almost always fall further and further behind. Often they drop out.

Tracking also can have the effect of racially segregating a school. African Americans, Latinos and recent immigrants are forced into the bottom tracks solely because of low test scores. In many districts, advanced classes are overwhelmingly filled with English-speaking white children.

BEST COPY AVAILABLE

Students who score the lowest may be labelled something like "educable mentally retarded" (EMR). In California in the 1970's, African Americans were less than 10% of the total school enrollment but more than 25% of all students in EMR classrooms. After examining all the evidence, a federal court found that the exams were biased. The judge banned the use of IQ tests for placing African American children in EMR classes in California.

Many schools and districts are beginning to change their placement policies to reduce or stop tracking. If tracking stops, the testing that is done to justify tracking can also stop.

Grade Promotion

Some school districts require a child to pass a standardized test before being promoted to the next grade. Students who do not pass the test are kept in the same grade: they are "retained" or "flunked." This is supposed to give these children a second chance to learn the material.

However, studies show that retention causes more problems than it solves. Flunking a child generally does not lead to improved school work. In fact, a student who is promoted usually does better than a similar child who is held back. Children who are retained often suffer psychological damage and a loss of self-esteem. Students who have flunked a grade are also much more likely to drop out of school. Fortunately, more schools now provide extra help to students without holding them back.

Despite all the talk about using tests to improve educational quality, the primary purpose of testing is still to sort children into groups. Heavy use of tests pushes schools to use practices — sorting, tracking, and flunking — that are more harmful than helpful.

"After 20 years of schooling, your aptitude test shows that you're skilled at just one thing — taking tests."

A Brief History of Testing

The primary early use of testing in education was to sort people into categories. The first standardized test, the IQ test (short for "Intelligence Quotient"), was developed in France a hundred years ago. It was designed to discover school children who might need special help and was also used to assign some immigrant children from Africa and Asia to special programs while ensuring that other schools remained white.

In the early 1900's, use of the IQ test was encouraged in the U.S. by followers of "eugenics" — people who believed intelligence was inherited and that certain groups of people (those from Northwestern Europe) were naturally smarter than others.

During World War I, supporters of eugenics convinced the government to experiment with using IQ tests to assign jobs to soldiers. These tests were unusual and their results were just as strange. Thousands of people were labelled "morons" or worse simply because they couldn't answer questions like:

- **The number of a Kaffir's legs is: 2, 4, 6, 8**
- **Christy Mathewson is famous as a: writer, artist, baseball player, comedian**
- **Karo is a: patent medicine, disinfectant, toothpaste, food product**
- **Kelvin was most famous in: politics, war, science, literature**
- **A puck is used in: tennis, football, hockey, golf**

The tests had no relationship to whether a person could actually perform certain tasks. Some researchers say they were rarely used in World War I to assign soldiers to jobs.

The testers were not stopped. They pushed to use similar exams to determine which immigrants should be allowed to enter the U.S. Once again, whole groups of people — such as Southern and Eastern Europeans, Africans, Asians, Latinos and Jews — were labelled "feeble-minded" and inferior. Immigration laws were changed to keep these people out of the U.S.

Testmakers continued to urge broader use of their products. In the 1920's, multiple-choice achievement tests, based on the same ideas as IQ tests, were developed. In that decade, IQ and achievement tests began to be commonly used to sort children into groups for instruction. Though many schools administered both kinds of tests, a student typically would take only a few standardized tests over 12 years of schooling.

In the 1960's, the development of computers that could score thousands of exams by machine made standardized exams cheap to score. At almost the same time, public officials were looking for quick and easy ways to improve education and win the space race with Russia. During this time, testing companies began to have a growing influence over education. Their salesmen convinced many school districts that multiple-choice achievement tests were the best way to rate student performance. They also persuaded the federal government that these tests were the best way to measure the progress of students in special programs like Chapter I, a federally-funded program providing extra help to low-income students who are falling behind academically.

In the 1980's, several major reports concluded that many children were not learning much in the nation's schools. Once again, increased concern about educational quality led to more standardized testing. This time, test scores were used to evaluate programs, schools and entire districts. Many states began requiring students to pass "minimum competency" tests for promotion or graduation. Even teachers were forced to pass standardized tests.

Today, more than 100 million standardized tests are given each year in the nation's classrooms. The typical student has to take between two and four each year.

BEST COPY AVAILABLE

III. WHAT ARE STANDARDIZED TESTS?

Tests are called "standardized" when the same questions are given to large groups of students under similar conditions. Most standardized tests are multiple-choice and are scored by a computer which has been programmed to give credit for only one "right" answer for each question.

You probably took tests like these when you were in school. They were the tests which made you fill in the blank or darken little circles or boxes with a "Number 2" pencil.

Some states and large school districts develop their own standardized tests, but most are made by private companies. Some test manufacturers are small firms. Many, however, are big businesses. In fact, the public school testing field is dominated by three huge companies — MacMillan/McGraw-Hill, Harcourt Brace Jovanovich, and Houghton Mifflin — that earn tens of millions of dollars each year selling their tests. Most college admissions tests are made by the Educational Testing Service and the American College Testing Program, both giant corporations.

Just because a test is said to be "standardized" does not mean it is accurate or meaningful. Test companies claim their exams are scientifically designed to measure a child's ability or achievement. Most measure little more than how well a student can quickly select the answer the test-maker wants from a list of four or five choices.

Standardized multiple-choice tests are sometimes called "objective" tests. This is because the test is scored by a machine or by using a scoring sheet — no individual decides a child's score. However, these tests are not objective in more important ways. Human beings still choose which questions to ask and which words to use in asking them. Decisions about what the correct answer is, what the incorrect choices are, and what a passing score will be are still made by test-makers, not machines. School officials still have to determine what to do with the results of a test. All these vital decisions are made by people who are, like everyone else, subjective.

Two Types of Standardized Tests

Standardized tests come in two major varieties: norm-referenced and criterion-referenced. It's important to know a little about how each type of test works because their results have very different meanings.

Norm-Referenced Tests (NRT)

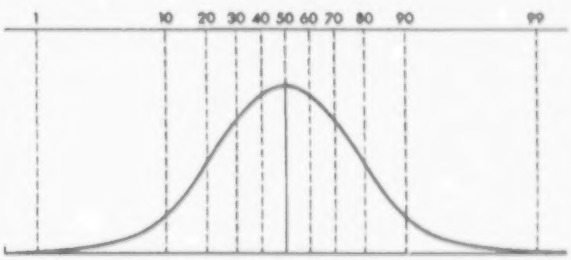
"Norm-referenced tests" (NRT) compare each child's score against a group of students who have already taken the same exam. The score reports whether a child is better than average, average, or worse than average, as compared with these other students.

Test-makers construct an NRT so that half the students taking it should end up below the middle score and half above. To expect all the students to be above average — above the 50th percentile — is like expecting all the teams in a baseball league to be above average. Just as in baseball, for each "winner" there must be a "loser" with norm-referenced tests.

Test-makers also have decided that most students will score near the middle and few students will get very high or very low scores. They assume that test results should end up arranged according to a bell-shaped, or "normal," curve (see chart, page 9).

However, the knowledge and abilities people have do not necessarily fit this model. The "normal" curve exists for the convenience of test-makers, not because it reflects reality.

Percentile Ranks in a "Normal Curve"



The numbers across the top indicate the possible rank. For example, all students to the left of the 10 scored below the 10th percentile. Most students score in the middle range, where the curve is highest.

To produce the curve test-makers want, the test has to be "normed." Many questions are administered to a small group of students (a "norming sample"). Testmakers then choose which questions ("items") will make up the test. They keep only those items that produce results that can create a curve. Some very tricky or confusing items are kept, but those most students can get right are usually thrown out. In making an NRT, it is often more important to choose questions that sort people out than it is to make sure that the content covered by the test is adequate.

Sorting test scores along the curve allows scores to be compared. For example, Sally is said to have scored at the 75th percentile. This means that one-quarter (25%) of all students at her grade level who were in the test norming sample scored the same as or higher than she did, while three-quarters (75%) scored lower than she did.

Criterion-Referenced Tests (CRT)

"Criterion-referenced tests" (CRT) do not compare students with one another. They simply measure whether a child has mastered a specific learning objective. The multiple-choice motor vehicle test many people take to get a driver's license is an example of a criterion-referenced test. It's possible for everyone to earn a passing score on this exam if they know about traffic signs and driving rules.

But suppose that the driver's license test asked questions about how a car engine works or was written in a language the driver didn't know. Then it wouldn't tell us much about whether a person knows the rules of driving, even though it was criterion-referenced. A test is not better just because it is criterion-referenced.

Unfortunately most CRTs are multiple-choice and essentially the same as norm-referenced tests. They can cause the same damage to children and education that norm-referenced tests cause.

Standardized tests do not have to be multiple-choice or short-answer. They can have essay questions or ask students to perform a task or make a product. For example, science students could be asked to perform and explain a lab experiment, rather than just answer word problems. New York State has a fourth-grade standardized lab test that works this way. Connecticut is developing extensive performance-based high school science assessments. (For better ways to evaluate students, see Section V.)

WHAT TEST SCORES MEAN

- A ***Raw Score*** is the total number of questions a child got right on a test. On an exam with 20 items, raw scores could range from 0 to 20. A student's raw score does not tell you much unless you know how many questions there were on the test and how difficult they were.

- A ***Percent Correct*** score is created by dividing the number of right answers by the total number of questions. For example, a child who gave the wanted answers to 10 out of 20 problems would have 50 percent (50%) correct.

- A ***Percentile Score*** or ***Rank*** tells you how a child's test score compares with the students on whom the test was normed. This is not the same as the percent of questions he or she got right. A percentile or rank of 50 means the student's score was better than the scores of half the students in the norming sample (see chart, page 10).

- A ***Standard Score*** is designed to report how much a student's score differs from the average score. For example, on an exam with an average score of 100, like most IQ tests, a score of 110 is as much above average as a 90 is below average.

- A ***Stanine Score***, short for standard nine, is one type of standard score. Stanines divide test results into nine groups, with 1 given to the lowest scores and 9 the highest. Scores of 4, 5 and 6 are considered average.

- A ***Normal Curve Equivalent*** (NCE) is another type of standard score, with an average of 50 and a range of 1 to 99. On any test, each grade has its own NCE. If a child had an NCE of 45 in second grade and 55 in third grade, he or she has made progress according to the test, since his or her scores rose from below average to above average. Chapter I score reporting is done in NCE's.

- A ***Grade Equivalent Score*** tries to compare your child's test results to the grade and month of the school year when that score would be average. A student who does as well as the typical test-taker in the fifth month of the second grade would get a grade equivalent score of 2.5. Using this method, half the students who take the test nationally should score below grade level and half above. Grade equivalent scores are misleading for students who do very well or poorly on the test. For example, a student in the fifth month of the second grade who scored a 4.9 is not doing fourth grade work and is almost certainly not able to do fourth grade work. He or she simply is doing very well, as compared with other second graders and as measured by the test.

All tests have **measurement error**. Consider a student who scores a 57 on a typical norm-referenced achievement test with a 0–100 score range and a normal curve. Two-thirds of the time, his or her score would fall between 50 and 64 — the **confidence band** or **score range**. The confidence band is supposed to tell you how much your child's score is likely to change if he or she took the same test again. In fact, the scores can vary even more than the confidence band says is likely. In this example, **one-third of the time, the score would be below 50 or above 64.** When a student takes a test, there is no way to know if, on that day, the score is inside or outside the confidence band.

IV. WHAT'S WRONG WITH STANDARDIZED TESTS?

Defenders of standardized tests say the problem isn't the tests but how they are used. Certainly test scores are often misused. But even when they are used in the way testmakers say they should be, they can still cause great harm. That's because standardized tests are fundamentally flawed. This section will show you how and why.

Test Results Are Often Wrong

Testers claim that standardized tests accurately and usefully measure a child's achievement, abilities or skills. In reality, current standardized tests are often neither *reliable* nor *valid*.

A test is completely *reliable* if you would get exactly the same results if you administered it to the same student again. For most tests, results can change a lot depending on testing conditions, how the student is feeling that day, the way directions are given, or whether the student is coached. As a result, many individuals' scores are wrong (see "Measurement Error," page 10). Young children are especially affected by these factors, so tests are much less reliable for them.

Because of problems with reliability (the "measurement error" of a test), almost all experts agree: test scores should never be the sole or primary basis for making a decision about a child.

Flawed Test Items

The following items come from achievement tests that are still in use. These flawed items also show part of the problem with using multiple choice items.

"The meat that was in the _____ burned."

- **hot**
- **oven**
- **cooking**
- **dinner**

(Metropolitan Achievement Test, 6th edition, Grade 2, 1985)

The wanted answer is oven. Why is dinner incorrect?

Find the picture of someone who most probably comes from another country.

(Comprehensive Tests of Basic Skills, 4th Edition, grade 1.0–2.2, 1989)

Why should any of these children "probably come from another country"?

BEST COPY AVAILABLE

How's Your IQ?

"How tall is the average American man?"
"Any answer from 5'7" to 5'11" is correct.

Children often answer with the height of their own fathers. Asians and Latin Americans are often shorter, while some other men are taller. How does a child's answer to this question show "intelligence" or lack of it?

"What is the thing to do when you cut your finger?"
2 point response: "Put a bandaid on it."
1 point response: "Go to the doctor (hospital)...Get it stitched up."
0 point response: "Suck blood...Don't panic."

Minority children usually perform poorly on this item. A sociologist from Baltimore, Maryland, asked inner-city youths why they answered the question the way they did. Many of these children answered "go to a hospital" because they thought that "cut" meant a big cut. When the youths were told that "cut" meant a little cut, almost all then responded, "Put a bandaid on it."

Both questions are from the Wechsler Intelligence Scales for Children - Revised (WISC-R) I.Q. test, currently given in many schools.

Validity is equally important. It refers to whether a test truly measures what it claims to measure. The information test-makers provide about validity is often incomplete and misleading. Just because the test maker says a test has been validated does not mean it really is valid or that it is good enough to be used to make decisions about children.

A test which is sold as valid for measuring one thing may really measure something else. For example, a "writing" test that asks students to mark the wrong word in a sentence does not measure writing, it actually measures proof-reading. In addition, tests validated for some particular use, such as screening, may be invalid if used for a different purpose, such as pupil placement. Such test misuse is very common across the U.S.

A supposedly valid exam also may be a poor measure of the subject it is supposed to test. For example, a biology test that mostly asks students to define words is a poor sample of what the subject of biology is really about.

A test that is supposed to be a valid predictor of how well a student will do may only be accurate because of how schools use the test results. For example, students with low test scores may be placed in a "slow" track. The fact that these children don't do well on later tests may be due to their poor education in the low track, not because they lack ability.

Because of inadequate reliability and validity, no test score is accurate enough to be used as the sole or primary basis for making decisions such as student placement or grade promotion. Unfortunately, schools often make these decisions based on test scores.

Many Tests Are Biased

Every test reflects the background of the people who design it. Most test-makers are white, upper to middle class professionals. Their tests reflect their own experiences, not those of many of the children taking the test.

This doesn't mean test-makers are racist, sexist, or elitist. They're simply human and limited by the fact that no one test can account for all the different cultures that make up our diverse society.

Any item written in a way that favors one group of students will put others at a disadvantage. Questions on topics of particular interest to African Americans, Latinos, or Asians will increase their scores, but such items are rarely included. Saying children have "low ability" because they don't have the background or experiences of white middle class children is clearly unfair.

Much test bias comes from the language on the exams. Tests are often written with very complex grammar and unusual words. This "super-standard English" may help increase differences among scores, but it is not fair to many test-takers.

S.A.T. QUESTION #1:
"IF TAD IS SUPPOSED TO MEET BUFFY AT THE COUNTRY CLUB, BUT HIS UOLUO BREAKS DOWN, WILL BUFFY PLAY TENNIS WITHOUT HIM?"
WHAT THE @#%& IS A "TAD"??

Tests can also be sexually biased. Among very young children, some tests are biased against boys. For teenagers the bias is in the other direction; teenage girls score lower on standardized college admissions tests, even though they earn higher grades than boys in high school and college.

Test-makers say low-income and minority-group children get lower scores because their schooling is usually not as good. This is partly true, but the problem is not solved by biased, unreliable, and invalid tests that lead to tracking.

In fact, heavy use of standardized testing usually makes matters worse for students who already attend inferior schools.

Tests Are Often Improperly Administered

Untrained test administrators, awkward seating arrangements, and overheated, chilly, or noisy rooms can all produce lower scores. The impact varies among different types of students. Young children are affected the most. Unfamiliar test examiners have been shown to hurt the scores of low-income and minority test-takers, but not the scores of middle income or majority students. *These problems make test results even more unreliable, but they are often ignored when looking at children's test scores.*

Tests and Scores Can Be Wrong

Like other human beings, test-writers make mistakes. Sometimes a question will have more than one "correct" answer. Other times no choice may be totally right. The answer the test-maker thinks is correct may not even be the best answer. Still, a child who marks two choices or leaves the question blank will lose points.

BEST COPY AVAILABLE

What Rights Do Parents Have About Testing Their Children?

Unfortunately, as a parent you have little control over your child's testing. Under U.S. law, you are allowed to see your child's test scores, along with all other records schools keep about your child. You have the right to attach a note to any test information protesting anything you think is wrong.

As a parent, you also have the right to insist that your child be tested in his or her first language if that language is not English. If your child has a learning disability, you may have the right to insist on special accommodations, such as having the test read aloud or using a large-print version. Aside from these, the law allows parents few rights over testing.

Parents *should have* many more rights. You should demand that your child's school or district adopt the following parent and student rights if they have not done so already:

- As a parent, you *should* be able to look at a test and review the questions, the wanted answers, and your child's answers with a teacher or administrator. A few states passed laws requiring manufacturers of college admissions tests to give copies of past test questions and correct answers to any student who wants them. Many errors in the tests have been discovered through this "Truth in Testing" law and thousands of scores have been raised. However, this protection does not apply to public school tests.
- As a parent, you *should* have the right to have tests and test scores explained in terms you understand. Schools should tell you when testing is going to happen, what kinds of testing is done, and how the results will be used. They should also give you your child's scores and answer any questions you have.
- As a parent, you *should* be able to challenge test scores that don't seem right. Remember, you can attach a note to the test score explaining why the score may be wrong.

If you oppose testing, you can try to get your children out of it. Unfortunately, if all the other children are being tested, you cannot legally refuse to have your daughter or son tested. You may have the right to refuse testing if most other children are not being tested. For any special testing, you should insist that the school explain the test and how the results might be used by the school, before your give your approval.

As a parent, you also have few rights over how tests are used to place children in school. Federal law does require a school to get your approval before your child can be placed in anything other than the regular program. Most tracking is considered part of the regular program. Assignment to *special education classes* or *bilingual education*, however, is not. In one case, the New York State Education Department ruled that parents can reject a school's decision to place a child in a *transitional* kindergarten or first grade program.

If you think your child has been incorrectly placed or tracked due to test scores or any other reason, you should challenge the decision. You should carefully investigate any special placement the school proposes for your child. Decisions should *never* be made primarily on the basis of test scores — no test is good enough. Broader information about the student in the school, at home and in the community — information from parents, teachers, and others — should be the basis for making educational decisions.

As a parent, you should insist on meetings where you can find out about the tests your school uses. At the meetings, administrators can inform you about testing policies, and you can ask questions to ensure that tests are not misused. You and sympathetic school personnel can also use these forums to persuade other parents and school officials to find ways to change assessment policies. Finally, the meetings can be an opportunity to learn about new methods of evaluating student and school progress.

The machines that grade answer sheets can also make errors. They can be thrown off by a stray mark on the paper, electrical problems, or other causes. Finding the right bubble in which to mark the answer is particularly confusing for young children. *As a result, the scores reported for one student — or an entire class — can be wrong.*

Standardized Tests Penalize Imagination and Critical Thinking

A child who thinks creatively, carefully, or deeply may have trouble with multiple-choice tests. These exams reward students who quickly select the "best" answer without thinking too hard. A child who thinks about why one choice is better than another may not be able to complete the test or may be forced to rush at the end. In either case, the result will be a lower score.

Some children have a hard time thinking the way test-makers want them to. They simply don't do well at the multiple-choice game. Other students get nervous because of the time pressure. Some, particularly young children, may even get physically sick.

Test-taking skills are the opposite of the good study habits our schools should teach. When the goal of schooling becomes high test scores, imagination and critical thinking get driven out of the classroom.

Tests Don't Measure Important Thinking Skills

Standardized multiple-choice tests don't deal with reasoning, problem-solving, or idea formation — the higher order thinking skills. Instead they focus on simpler tasks like recognizing facts. These are not the skills needed to produce scientists, artists, teachers, or the other adults our society needs.

Multiple-choice tests assume that every problem has only one correct answer. The truth is, many problems have more than one solution. In the real world, often the first problem is to figure out what the problem is. Multiple-choice tests tell you what the problem is without making you figure it out. In no subject area can a multiple-choice test adequately evaluate higher order thinking or the ability to use knowledge to solve real problems. For example, a mechanic who cannot pass a multiple-choice test on mechanics still may be able to fix your car. Nor can any standardized test measure "people skills," imagination, creativity, determination, or leadership.

Even though your child may not do well on standardized tests, he or she may still do well in school and in later life. Most of what our children need to know cannot be measured by a standardized test.

Tests Don't Reflect How People Actually Learn

The ideas behind achievement tests were developed by nineteenth century psychologists. They thought that learning is a process in which a person passively absorbs separate bits of information until she or he has enough bits to assemble a whole idea.

Now psychologists know that learning is active. People — including young children — learn by connecting what they already know with what they are trying to learn. If they cannot actively make meaning out of what they are being taught, they are unlikely to learn or remember. Even learning basic

skills — like speaking and reading — involves complex thinking. Students can and should strengthen basic and higher order abilities at the same time.

While our understanding of the brain and how people learn has progressed enormously, standardized tests have remained the same. Today, most educators — including even the test-makers — agree that schools should not make curriculum and instruction look like the tests. When instruction is controlled by the tests, students become passive and unthinking. In the end, they forget most of what they "learned" because it had no meaning for them. Not only do students fail to acquire higher order skills, they don't even learn basic skills very well.

Using tests and evaluation methods that encourage students to think can also help teachers improve their teaching. These better ways to evaluate are discussed in Section V.

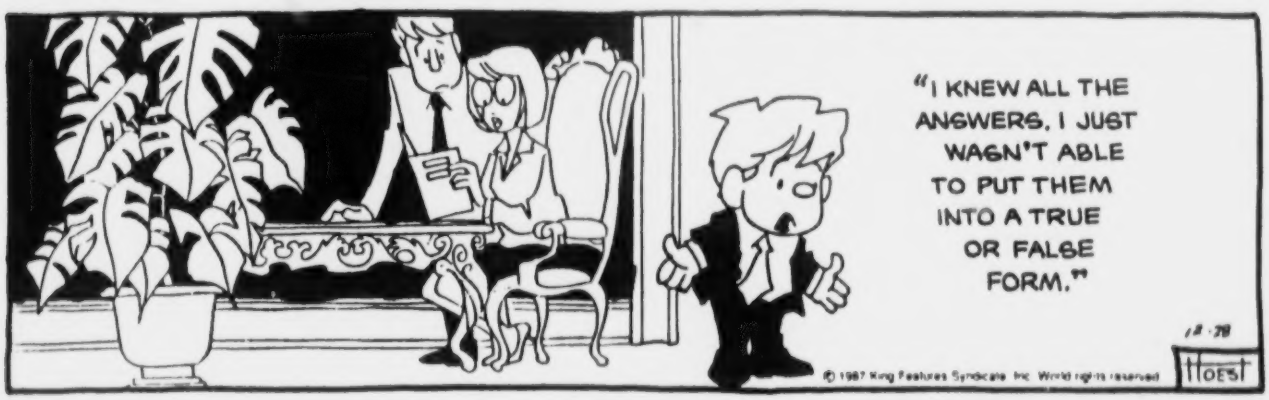
Test Scores May Not Reflect What A Child Has Actually Learned

A standardized test is designed to be given to children all across the country. Its content may have little to do with what is taught in any particular school. Not all classes cover material in the same order. If children are tested on material their teacher has not taught, their scores will suffer.

Testing Hurts All of Education

Multiple-choice testing has taken over education. Despite the flaws of the tests and the harm they cause to millions of children, their use has come to control what public schools cover and what teachers teach. Increasingly, however, parents, teachers and many others have come to see that this is a serious mistake. Depending heavily on testing has not made education better. In fact, it prevents needed improvement and gives too much power to testing companies over which parents, teachers, and communities have no control.

WHAT A GUY! / Bill Hoest



Reprinted with special permission of King Features Syndicate, Inc.

BEST COPY AVAILABLE

Standardized Tests Damage School Quality

Increasingly, students' test scores are used to rate teachers, schools, and entire districts. Across the country, newspapers report average scores as if they were an accurate measure of school quality. In some states, the level of school funding depends on test results.

All these factors make for "high stakes" testing. When school officials and teachers feel pressure to improve test scores, they start to "teach to the test." This means that classrooms are turned into test coaching centers.

Some schools start in September to prepare students for tests given in April and May. Many schools start coaching in January. The result may be higher test scores, but all this preparation time robs children of important learning.

Teaching to the test ignores almost everything not included in standardized multiple-choice exams. If the reading test only asks questions about short paragraphs, that is what teachers will focus on. Novels, essays, and poems often are not studied. Many times, students in such classes don't read books, work on projects, solve real-world problems, or come to understand science or history. Instead, they simply memorize isolated facts and do worksheets that look just like tests. This happens most in low tracks where students are likely to be from low-income families and minority-groups. Ironically, these children are often in low tracks because of their test scores.

Textbooks are also "dummied down" so their form and content match the tests. Often they are made by the same companies that make the tests and are designed to help raise test scores. Some texts even include multiple-choice tests at the end of each chapter. As a result, students become less interested in reading and seldom remember what they read. Bored and frustrated, these students are more likely to drop out of school.

Supporters claim that focusing on testing leads to improvement in basic skills. They point to rising test scores on achievement tests as evidence. Unfortunately, rising test scores do not mean that more learning has taken place or that the quality of education has improved. Test coaching can increase scores without increasing real knowledge.

Despite a huge increase in standardized, multiple-choice testing in the 1980's, National Assessment of Educational Progress studies show that overall achievement has increased very slightly or not at all. In 1988, children in grade four — the students who have been most affected by a test-controlled curriculum — actually read less well than fourth graders read in 1980. All the evidence shows that students are not developing higher order skills the way they should. *When tests control the schools, test scores may go up, but educational quality usually goes down.*

Testing Damages Accountability

Those who favor testing say that it is needed for accountability. How else, they ask, can we know if our children are learning? Accountability is an important and legitimate concern of parents, citizens, and legislators. *Using a tool that is often wrong and damaging to education, however, is no way to provide genuine accountability.*

Because of the pressure to raise test scores, some schools even cheat. School officials have been caught giving students the answers, erasing errors, and destroying poor answer sheets. Students who test poorly may be encouraged to stay home on test days. Cheating scandals have arisen in California, New York, South Carolina, Virginia, and other states.

Cheating, coaching, and changing the curriculum to fit the test are some of the reasons every state reports test scores above the national average, even though that is logically impossible. If we cannot believe the test scores, then they must be useless for accountability.

Putting education under the control of distant corporations and taking control away from local parents, teachers, administrators, school boards, students, and other community members also harms accountability. The alternatives to standardized tests described in Section V are better ways to evaluate achievement that also increase accountability.

Is There Anything Good About Standardized Tests?

Not much. Sometimes a student from a poor or minority family who scores well on a test will get special attention — though his or her fellow students will not. Or schools with very low scores may be targeted for extra help.

Test scores have been used to draw attention to the lack of good schooling available for low-income and minority-group children. In some states, testing has been used to help ensure more students are properly served in bilingual programs.

Multiple-choice and short-answer test items could be used as a small part of a statewide testing program. If such a program sampled students and was not administered in every grade, it could provide useful accountability information without hurting individuals or the curriculum.

Overall, the use of standardized tests has done little to improve education for individual students or society. The few good uses are far outweighed by the harm to most children and schools.

V. BETTER WAYS TO EVALUATE STUDENTS

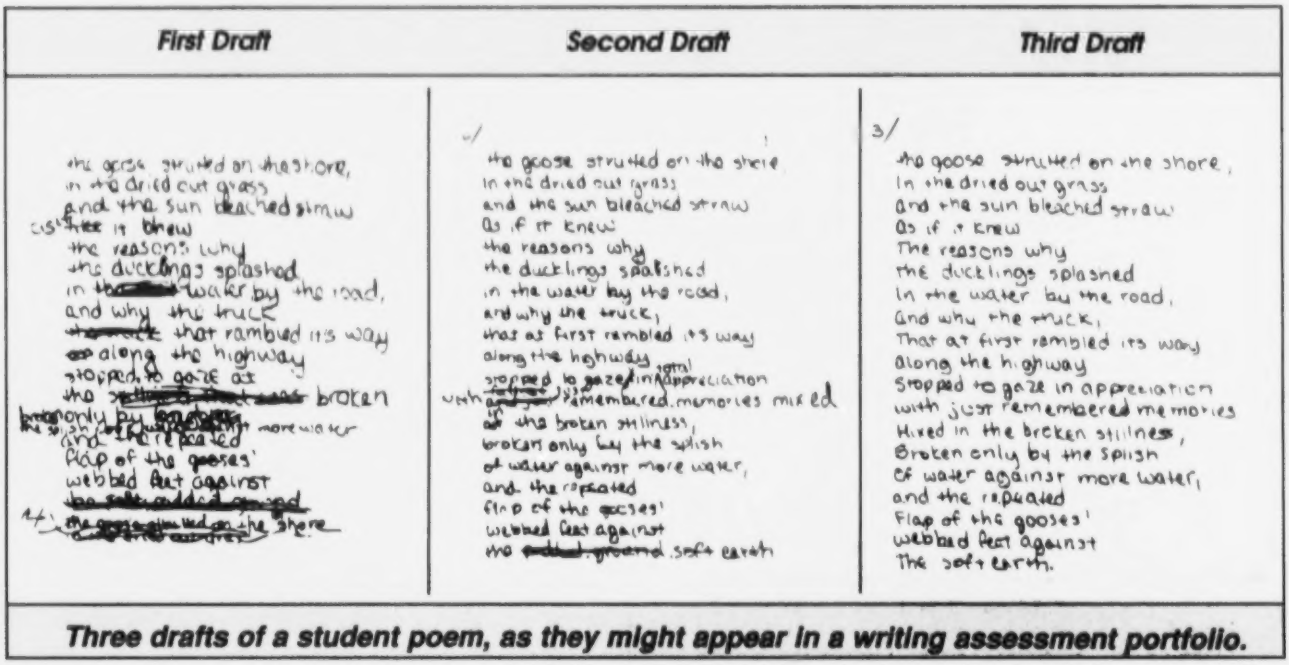
The most important purpose of evaluation should be to help learning. The New York State School Boards Association, for example, has said, "The primary goal of a testing policy should be to improve individual student learning." The second purpose of evaluation is to provide information to people outside the school — parents, school boards, legislators — about the quality of student learning.

Standardized multiple-choice tests fail both goals. Fortunately, better ways to evaluate and assess children now exist. These tools go by names like **documentation**, **performance-based assessment**, **authentic** or **appropriate evaluation**, and **direct measurement**.

All these alternatives to standardized exams are based on a simple notion: We should evaluate students as they actually perform tasks and demonstrate knowledge of the material we want schools to teach. For example, if we want students to be able to write clearly, schools should look at and grade the writing they do.

Observation and documentation are the most basic types of evaluation. The teacher watches each student's activity, takes notes, and combines this information with examples of the student's work. Keeping careful track of students' writing is an example of documentation. In reading, we can record what books a student has read, whether he or she understands them well enough to be able to retell them in his or her own words, and whether the student can make use of the experience or knowledge contained in the book.

A student's work samples can be kept in a **portfolio**. For writing, this folder may contain several drafts of a paper with com-



First Draft	Second Draft	Third Draft
the geese strutted on the shore, in the dried out grass and the sun bleached straw as if it knew the reasons why the ducklings splashed in the water by the road, and why the truck that rambled its way along the highway stopped to gaze at the broken only by the and the repeated flap of the geeses' webbed feet against the soft earth	the goose strutted on the shore in the dried out grass and the sun bleached straw As if it knew the reasons why the ducklings splashed in the water by the road, and why the truck, that at first rambled its way along the highway stopped to gaze in appreciation with just remembered memories mixed of the broken stillness, broken only by the splish of water against more water, and the repeated flap of the geeses' webbed feet against the soft earth	3/ the goose strutted on the shore, In the dried out grass And the sun bleached straw As if it knew The reasons why the ducklings splashed In the water by the road, And why the truck, That at first rambled its way Along the highway Stopped to gaze in appreciation with just remembered memories Mixed in the broken stillness, Broken only by the splish Of water against more water, and the repeated Flap of the geeses' webbed feet against The soft earth.

Three drafts of a student poem, as they might appear in a writing assessment portfolio.

From "Portfolio Assessment: Sampling Student Work," by Dennie Palmer Wolf. Educational Leadership, Vol. 46, No.7, April, 1989. Reprinted with permission of the Association for Supervision and Curriculum Development. © 1989 by ASCD. *All rights reserved.*

 BEST COPY AVAILABLE

ments from both teacher and student. Parents can look at the portfolio to see how well their son or daughter is doing. Outside examiners might look at a sample of portfolios to see whether an entire class is learning at a satisfactory level. The state of Vermont is beginning to use portfolio evaluation.

Documentation and portfolios can be used to help teachers become better instructors. This process can also be used to protect students from teacher bias. For example, other teachers, administrators, and trained community people can directly examine samples of student work, note the levels of student achievement in various subjects, and compare the teacher's evaluation with the student work they see in the portfolio. If the observers detect problems, they could interview teacher and students and sit in on some classes. This evaluation process can be the basis for working with teachers to improve their teaching and overcome any biases.

Teachers can summarize some of the information contained in portfolios by using **checklists** and **inventories**. These tools help record student progress in a systematic way.

Open-ended tests are another improvement. Students might be asked to solve a multi-step math problem from scratch. Instead of being able to "psyche out" the answer from the choices offered, they'll have to figure out what the problem really is, solve it, write down their work, and explain why they did what they did. Similarly, all students might be required to write a paper to show they can reason clearly and communicate well. For example, a class could be asked to write an essay designed to persuade the school board to stop unnecessary testing.

These tests can help teachers by showing how the student thinks and uses knowledge in different subject areas. They can be a method to check the results of portfolio work. They also can be useful to people outside the school. Arizona, California, Connecticut and Kentucky are among the states developing open-ended tests.

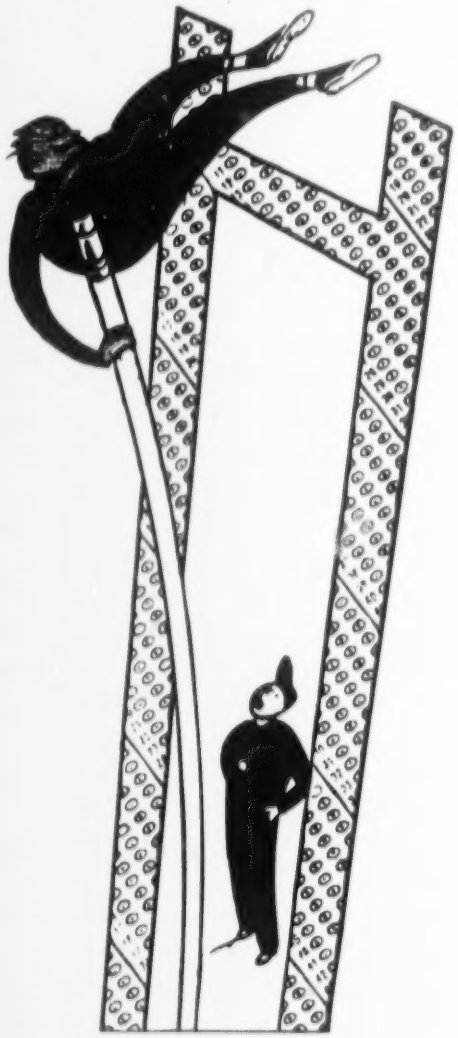


Illustration by Steve McCracken. Reprinted with permission.

BEST COPY AVAILABLE

An evaluation that is part documentation or portfolio and part test is a **performance**. Putting on a play, giving a musical recital, or competing in an athletic contest are all types of performance. So is making a **product** like a table in woodworking class or a drawing in art.

Performances and products can also be used for academic subjects. For example, students could do historical research, write up their findings (a product), and explain and answer questions about it (a performance). The end result can be scored, perhaps by a team of educators, students, parents, and other community members. Each student can get feedback about his or her work during the process and at the end. This already happens with projects for science fairs. Some schools even require students to create a major performance, product, or **exhibition** in order to graduate. All of this is a form of documentation as well as a type of test. Unlike taking multiple-choice tests, these can be worthwhile learning experiences for students.

Scoring Performance-Based Assessments

It's not hard to design scoring systems to replace standardized tests. In Olympic Games, for events like figure skating or gymnastics, judges give athletes a mark based on the quality and difficulty of their performance. The same thing can happen in the classroom. Teachers can be trained to give *ratings* based on a common scale. A score of "6" might be given for an excellent paper, a "5" for one that is very good, and so on down to a "1" for a very weak paper. In this way, students can be compared with each other and against a standard of excellence. We have learned from other countries which rely heavily on essays to evaluate students and from athletic competition and science fairs that trained raters generally agree about the right scores to give.

Similar systems can be used in other subject areas. Science assessment can be based on scoring research projects, laboratory experiments, essays, and open-ended questions. In the process of doing this work, students learn the important facts and rules that standardized tests now try to measure. More importantly, they learn how to think scientifically by working on real-world problems.

Although they all help teach students to think in more complex, meaningful ways, these testing alternatives are not perfect. Teachers can spend too much time coaching their students for the tasks on the tests or give scores based on their biases and not the students' performances. But outside evaluators and aware parents should be able to guard against these dangers. That's why it is very important to include checks and balances to ensure the test's accuracy and fairness. In addition, no test — no matter how good it is — should be used as the sole basis for making important educational decisions.

Properly designed and supervised, these new evaluation and testing systems will help improve teaching and learning. Classrooms can stop becoming test-coaching factories. Fewer students will be turned off by boring, test-driven material. Parents, educators, and the public will be better able to see how effective our schools are.

VI. WHAT YOU CAN DO

You can help reduce testing and bring authentic assessment to our schools by forming groups, coalitions, and campaigns that address these issues locally or statewide.

Campaign Goals

The more you learn about testing and alternatives, the more you can help. Information that you should find out includes:

- What tests are given in your local district and state?
- In which grades are students tested?
- How are test results used?

This knowledge will help you be a more effective advocate when added to the information contained in this guide. Do not allow lack of knowledge or experience to prevent you from getting started. You can learn the things you need to know.

Form a Group or Coalition

If you want to make changes, you must work with other people. By joining with like-minded individuals and groups, it is easier to get the word out about your cause and to get the work done.

If you are a teacher, you can begin by talking to other teachers in your school about how standardized testing affects teaching and students. You can bring the issue to your union or teacher group.

As a parent, you can talk to friends who are parents of school-aged children, or you can bring the issue to your local parent group.

If you are an administrator, you can examine how your school or district uses tests. Principals can work to lower the pressure of testing and help introduce alternatives. Superintendents can often reduce testing themselves, investigate alternatives, and help teachers to begin implementing them.

Teachers, parents, and administrators should try to discuss these issues with each other. If you can include different kinds of

TIGER by Bud Blake

Reprinted with special permission of King Features Syndicate, Inc.

BEST COPY AVAILABLE

people in your group or coalition, you will increase your chances of success.

It is not essential to begin with a large group. A small group of dedicated people can accomplish a lot.

In Denver City, Texas, two parents of elementary school children were the driving force behind Kids, Education and You (KEY). Their efforts to stop the practice of holding back children in pre-kindergarten and kindergarten, often on the basis of test scores, led to local changes and then helped persuade the Texas Board of Education to ban this practice. KEY's work was influential in getting the Board to rule that transitional classes between kindergarten and first grade are a form of holding children back. The Committee for Effective School Evaluation (CESE), an Arizona group that persuaded the state to drop standardized test requirements in grades 1 and 12, was led by just two people — a teacher and a parent.

Individuals or groups that may support your efforts include: parents, teachers, school administrators, school board members, local businesses, city or state legislators, and civil rights, community and educational organizations. Plan how to reach and persuade these kinds of people. Figure out who is most influential with decision-makers such as school boards. These are people and groups you will especially need to convince.

Establish Clear Goals

In forming a group or coalition, you must agree on an overall goal or purpose. In New York, for example, the people who formed the New York Campaign for Genuine Accountability in Education wrote a *Statement to Urge a Ban on Standardized Testing of Young Children*. Agreement on a goal in the beginning provides a continuing focus as the activities and number of people who join you increase. (See pages 27-28 for FairTest's national "Statement on Genuine Accountability in Education," signed by more than 75 organizations.)

You can develop short-term, medium-term, and long-term goals. A short-term goal might be to meet with the school board of your district to present your position and find out where the board stands. A medium-term goal might be to reduce or eliminate standardized testing in the early grades. A long-term goal might be to change state law regarding testing.

Each state/district/school is different. Goals, therefore, should be developed based on the particular needs of your community or state. The New York Campaign, for example, is focusing on the testing of young children because most districts test in the early grades even though the state does not mandate testing until grade 3.

Make use of experiences and resources already available. People involved in planning the New York Campaign learned from a North Carolina campaign which succeeded in stopping testing in grades 1 and 2. They also gained knowledge from the experiences of those who had worked in New York over the years to change the ways that college admissions tests are used.

In choosing your goals, especially short-term goals, select ones that you are likely to accomplish: "Nothing succeeds like success." Be sure, however, that the goals are important enough to make the effort worthwhile.

Develop a Strategy — Get People Involved

A strategy is a plan for reaching your goal. In it, you include all the major steps you must take in order to win. For example, the Atlantic Center for Research in Education's (ACRE) goal was to abolish state-mandated testing in grades 1 and 2 in North Carolina. This required passing state legislation. To win, ACRE, had to document problems, attend meetings, help organize teachers of young children, and persuade first some, then many, legislators to support them.

Educating key individuals, groups, policy-makers, and the general public is always necessary for achieving success. This can be done in a number of ways: brochures, public presentations, fact sheets, letters-to-the-editor, meetings, public service announcements, articles in organizational newsletters, workshops, leaflets, radio and television appearances such as talk shows, and videotapes all provide ways to get information out to the public. Distribution of this booklet will aid your public education work. Use the form on page 29 to order copies.

It is also important to know which materials or arguments to use with which people. ACRE found that their most persuasive arguments to legislators were: 1) young children are poor test takers, so test results are not accurate; and 2) standardized tests ignored very important educational goals for young children which could not be measured with paper and pencil tests. In your area, your best arguments may be different.

Ultimately, testing reform activists must convince public officials. You can change public policy by informing and pressuring these decision-makers at the local or state level. Learn how decisions are made. Find supportive groups that know the local or state decision-making process and have them help you.

Be Flexible

Preparation for problems that might arise should be a part of choosing goals and planning. Occasionally, however, something unexpected happens or new information becomes available which could affect your original plans.

Compromise may be appropriate. For instance, the Arizona group accepted a plan for the state to stop grade 1 and 12 testing even though their original goal was to stop testing in grades 1, 2, and 12. They had a choice between this or no change at all from the legislature.

If your strategy is not working, a completely different plan may be called for. You may need to make adjustments along the way. In any case, flexibility makes adjustments easier and allows you to make changes in your plans if necessary.

Be Prepared for the Long Haul

Systems that have been in existence for a long period of time, such as the public education bureaucracy, are not changed easily or overnight. By realizing that it could take several years to see final results from your work, you avoid the pitfalls which result from unrealistic expectations.

If you want information, materials (e.g. fact sheets), or assistance in getting started, contact the FairTest office at 342 Broadway, Cambridge, MA 02139; (617) 864-4810.

APPENDIX: Resources You Can Use

Bibliography

From FairTest, 342 Broadway, Cambridge, MA 02139.

- *Fallout from the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools* (3rd Edition, 1990). Contains the information and references to back up this guide. $8.95.
- *FairTest Examiner.* Quarterly newsletter on testing reform. $15/yr individuals; $25/yr institutions.
- *Standing Up to the SAT* (1989). College admissions testing reform and test coaching in one book. $6.95.

FairTest also has an annotated bibliography on testing in the public schools (included in *Fallout*) that covers all the different problems with testing. And FairTest has Fact Sheets on University Admissions Testing and K-12 Testing. (Contact FairTest.)

Other readings/resources

- *Network* (Back to School 1990): "High Scores vs. High Achievement: Concerns About Standardized Testing in America's School Improvement Movement." (National Committee for Citizens in Education, 10840 Little Patuxent Parkway, Suite 301, Columbia, MD 21044; 1-800-NET-WORK.)
- *Rethinking Schools* (Jan/Feb 1989): "Focus on Testing" section. (P.O. Box 93371, Milwaukee, WI 53202); $2.00.
- Southern Association on Children Under 6. *Developmentally Appropriate Assessment: A Position Paper.* (1990: SACUS, P.O. Box 5403, Brady Station, Little Rock, AR 72215; 1-501-663-0353.) 20¢.
- Kamii, C., ed. *Achievement Testing in the Early Grades: The Games Grown-Ups Play.* (1990: National Association for the Education of Young Children, 1834 Connecticut Avenue, N.W., Washington, D.C. 20009; 1-800-424-2460.) $8.00.
- *Educational Leadership*, April 1989: "Redirecting Assessment." 17 articles on authentic assessment. (Order Processing, ASCD, 1250 N. Pitt St., Alexandria, VA 22314.) $5.00.
- *Phi Delta Kappan*, May 1989: Special Section on Testing has article from FairTest and articles on authentic assessment. (P.O. Box 789, Bloomington, IN 47402.)
- Archbald, D. & F. Newman. *Beyond Standardized Testing: Assessing Authentic Academic Achievement in the Secondary School.* (1988: National Association of Secondary School Principals, 1904 Association Drive; Reston, VA 22091.) $8.00.

Standardized Tests and Our Children

Additional copies of this *Guide to Testing Reform* are available by mail from FairTest at the following prices:

Single copy:	$3.00
5 copies:	$10.00
10 copies:	$15.00
50 copies:	$50.00
100 copies:	$75.00

Use the order form on page 29.
For licensing arrangements to reprint large quantities, contact FairTest.

This Guide is also available in Spanish at the same prices.

APPENDIX: What Testing Terms Really Mean

Testing is full of terms that are not familiar to many people. Some words don't have the same meaning in testing as they do in regular use. In order to better understand these words when you talk with teachers, administrators, school boards, and others, here is a dictionary of basic testing terms.

Achievement Tests — claim to measure how much a student has learned in a class such as reading, spelling, science or math.

Aptitude Tests — try to predict a student's ability to do certain types of school work; sometimes called "cognitive abilities" tests; IQ tests may be used for this purpose.

Competency Tests — are designed to determine whether a student has met a minimum standard of skills and knowledge in a specific subject area.

Criterion-Referenced Tests — measure how well a student meets certain standards determined by the test-maker.

Cut Score — is the minimum level a test-taker must attain in order to "pass" a given exam.

Developmental Screening Tests — are given to identify students who may have problems which need attention from a specialist such as a doctor.

High Stakes Tests — are tests that are used to make important educational decisions and that educators and students view as important.

IQ (Intelligence Quotient) Tests — supposedly measure a person's native intelligence by assuming a single set of inherited abilities which can be measured by one test.

Items — are test questions or problems.

Measurement Error — is one way by which reliability or lack of reliability of a test is indicated; all tests have measurement error.

Norm-Referenced Tests — are designed so that the scores of all the students who take it fall along a "normal" or "bell-shaped" curve; most of the test-takers will be near the middle, few will be at the ends; half the test-takers will be "below average."

Reliability — estimates how closely a test would produce the same results if it were repeatedly administered to the same student under the same conditions.

Sampling —is a way to get information about a group by examining only some members of the group or by giving all members only small parts of the whole test.

School Readiness Tests — are used to determine whether a child is "prepared" to begin kindergarten or first grade.

True Score — the score a student should get if the test were completely reliable.

Validity — defines how accurately a test really measures what its promoters claim it measures; *predictive validity* measures how well the test predicts future school performance; *content validity* measures how well the test covers the subject content.

Statement of the Campaign for Genuine Accountability in Education

We, the members of the Campaign for Genuine Accountability in Education, believe that a heavy emphasis on standardized, multiple-choice tests as the primary measure for educational ability or achievement will undermine needed educational reforms in our nation. We have prepared this Statement in order to urge fundamental changes in assessment and evaluation.

It is crucial that political and education officials, at all levels, recognize that standardized, multiple-choice tests are not an adequate means to measure educational progress, nor are high test scores an appropriate educational goal. There are educationally sound means of evaluation which can provide a basis for improving teaching, informing the public, and measuring progress, thereby providing genuine accountability.

The states, local districts, and the federal government should embrace the goal of moving away from reliance on standardized multiple-choice test scores to use of alternative, authentic and appropriate methods of assessing educational performance and progress. In doing so:

- **Education and political leaders should call for the use of appropriate and authentic evaluation in classrooms and to assess aggregate student achievement and progress; they should call for their states and local districts and federally funded programs, such as Chapter I, to implement available alternative assessment measures and continue to develop new ones for use in evaluating students and programs.**

- **These leaders should not call for raising standardized test scores as a goal of educational reform; should not call for new or additional testing with standardized, multiple-choice exams; should not rely on existing standardized tests as indicators of educational progress; and should not depend on standardized tests to make decisions about program placements for students. These tests misdirect the curriculum, are inadequate educational measures, are not instructionally useful, are often too inaccurate to be the basis for making decisions about individuals, and can be replaced with authentic assessment methods.**

The alternatives we support go by various names: authentic, appropriate, performance-based, and direct. They are evaluations based on asking students to demonstrate what we want them to be able to know and do. For example, if students should write well, we should look at their writing. Appropriate assessments include teacher observations, performances, exhibitions, portfolios of student work, checklists, and open-ended questions that encourage real thought. These assessments provide instructionally useful information, can avoid the cultural and racial biases of many standardized tests, and can be put into a form that will provide useful data and genuine accountability to local, state, and national governments.

Political and educational leaders at all levels should set a time-table for phasing-out most uses of standardized, multiple-choice tests and replacing them with reliable, valid, and educationally appropriate alternatives. In the period before the implementation of alternatives, use of and reliance on standardized multiple-choice tests should be reduced as much and as quickly as possible.

List of Signers

"Statement of the Campaign for Genuine Accountability in Education"

Advocates for Children of New York
American Association of Colleges for Teacher Education
American Federation of Teachers
American Reading Council
Angelo N. Ancheta, Asian Law Alliance*
APPLE Corps., Inc.
Association for Supervision and Curriculum Development
Bank Street College of Education, New York
Rims Barber, Mississippi Human Services Agenda
Leonard Beckum, Dean, College of Education, CUNY
Association for Women in Science
Diana Caballero, Puerto Rican/Latino Education Roundtable, NYC*
Center for Law and Social Policy
Center for Women Policy Studies
Council for Basic Education
Robert W. Covert, Past President, American Evaluation Association*
Harold E. Dent, Vice-President, Psychological and Human Resources Consultants, Inc.*
Educational Law Center, Inc.
Pablo Eisenberg, President, Center for Community Change*
Federation of Organizations for Professional Women
Foxfire Fund, Inc.
Fund for the Feminist Majority
Howard Gardner, Project Zero, Harvard University*
Genesee Valley Developmental Learning Group, New York
Leslie A. Hart, Brain-Compatible Education Associates
La Donna Harris, Americans for Indian Opportunity*
Asa Hilliard, Professor of Education, Georgia State University*
Holt Associates/*Growing Without Schooling*
Institute for Learning and Teaching
International Reading Association
Kentucky Association on Children Under Six
KEY-Kids, Education and You
A. Gay Kingman, National Congress of American Indians*
Louisiana Association on Children Under Six
Massachusetts Advocacy Center
Matsushita Foundation, Inc.
Deborah Meier, Principal, Central Park East Secondary School, NYC*
Samuel J. Meisels, Professor of Education, U of Michigan*
Sara E. Melendez, Vice-Provost, University of Bridgeport*
Mexican American Women's National Association
National Association for the Advancement of Colored People
National Center for Fair & Open Testing (FairTest)
National Coalition of Advocates for Students
National Coalition of Title I/Chapter I Parents
National Council of Teachers of English
National Council for Social Studies
National Education Association
National Indian Youth Council, Inc.
National Organization for Women--New York City Chapter
National Parent Teacher Association
National School Boards Association
National Women's Law Center
National Women's Political Caucus
Fred M. Newmann, National Center on Effective Secondary Schools--University of Wisconsin*
New York Public Interest Research Group
NOW Legal Defense and Education Fund
Organization of Chinese American Women
Vito Perrone, Harvard University--Graduate School of Education*
Project Equality, Inc., Kansas City, MO
Puerto Rican Legal Defense and Education Fund, Inc.
Rochester (NY) Teachers Association (AFT)
Lori Rubenstein, Partnership for Democracy*
William Saunders, National Alliance of Black School Educators*
William V. Schipper, Executive Director, Natl. Assn. of State Directors of Special Education*
Hilton Smith, Coordinator, Foxfire Teacher Outreach Programs*
Southern Association on Children Under Six
Southern Christian Leadership Conference
Southern Regional Council, Inc.
Adria Steinberg, Harvard Education Letter*
United Federation of Teachers (New York City)
United States Student Association
Dorothy J. Watson, President, Whole Language Umbrella*
Paul Weckstein, Center for Law and Education*
Grant Wiggins, Center on Learning, Assessment and School Structures (CLASS)
Arthur E. Wise, NCATE*
Women's Research and Education Institute

***Organization for identification purposes only**

Many other individuals, particularly teachers, have endorsed the statement.

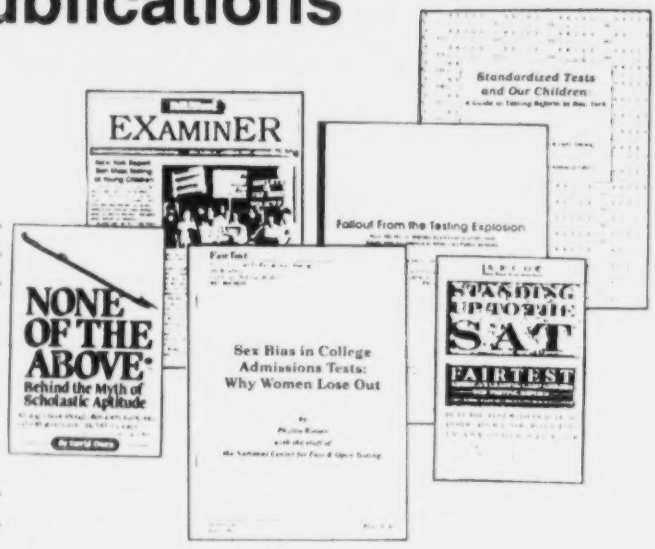
FairTest Publications

Standardized Tests and Our Children: A Guide to Testing Reform, by FairTest (32 pp.)... 1 for $3; 5 for $10; 10 for $15; 50 for $50; 100 for $75. English and Spanish editions can be combined for bulk prices.

English _____
Spanish _____
Total _____

Standardized Tests and Our Children: A Guide to Testing Reform in New York, by FairTest & NYPIRC (32 pp.)...same prices as above. Single copies from FairTest. Bulk copies available from: NYPIRC, 184 Washington Ave., Albany, NY 12210

English _____
Spanish _____
Total _____


Standardized Tests and Our Children
EXAMINER
Fallout From the Testing Explosion
NONE OF THE ABOVE
Behind the Myth of Scholastic Aptitude
Sex Bias in College Admissions Tests: Why Women Lose Out
SAT
FAIRTEST

FairTest Examiner, published quarterly...$15.00/year (individuals), $25.00/year (institutions) _____

Back issues of ***FairTest Examiner***, $4.00 each: Vol __ No. __ (All 13 for $40.00) _____

Standing Up to The SAT, by John Weiss, Barbara Beckwith and Bob Schaeffer (192 pp.)...$6.95 _____

Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools, by Noe Medina and Monty Neill(45 pp.)...$8.95 _____

Sex Bias in College Admissions Tests: Why Women Lose Out, by Phyllis Rosser and FairTest staff, 3rd ed. (60 pp.)...$7.95 _____

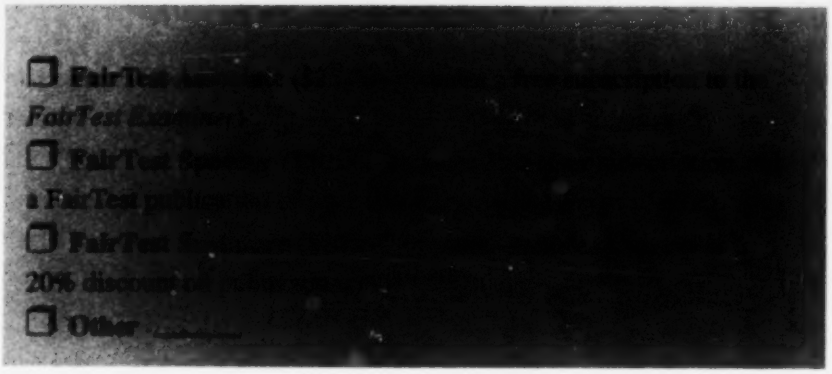
The Reign of ETS, by Allan Nairn and Ralph Nader (550 pp.)...$30.00 _____

Beyond Standardized Tests: Admissions Alternatives That Work, by Amy Allina with FairTest staff (18 pp.)...$5.50 _____

None of the Above, Behind the Myth of Scholastic Aptitude, by David Owen (327 pp.)...$5.95 _____

Contributions to FairTest

I agree there's no better way to be part of the testing reform movement than supporting FairTest! Enclosed is my tax-deductible contribution!*



Publications Total _____
Contributions Total _____
Check Total _____

Name _____
Address _____
City _____ State _____ Zip _____

Please make checks payable to FairTest and mail to:
FairTest, 342 Broadway, Cambridge, MA 02139

BEST COPY AVAILABLE

*Contributions to FairTest are fully tax deductible, except for the value of publications received.

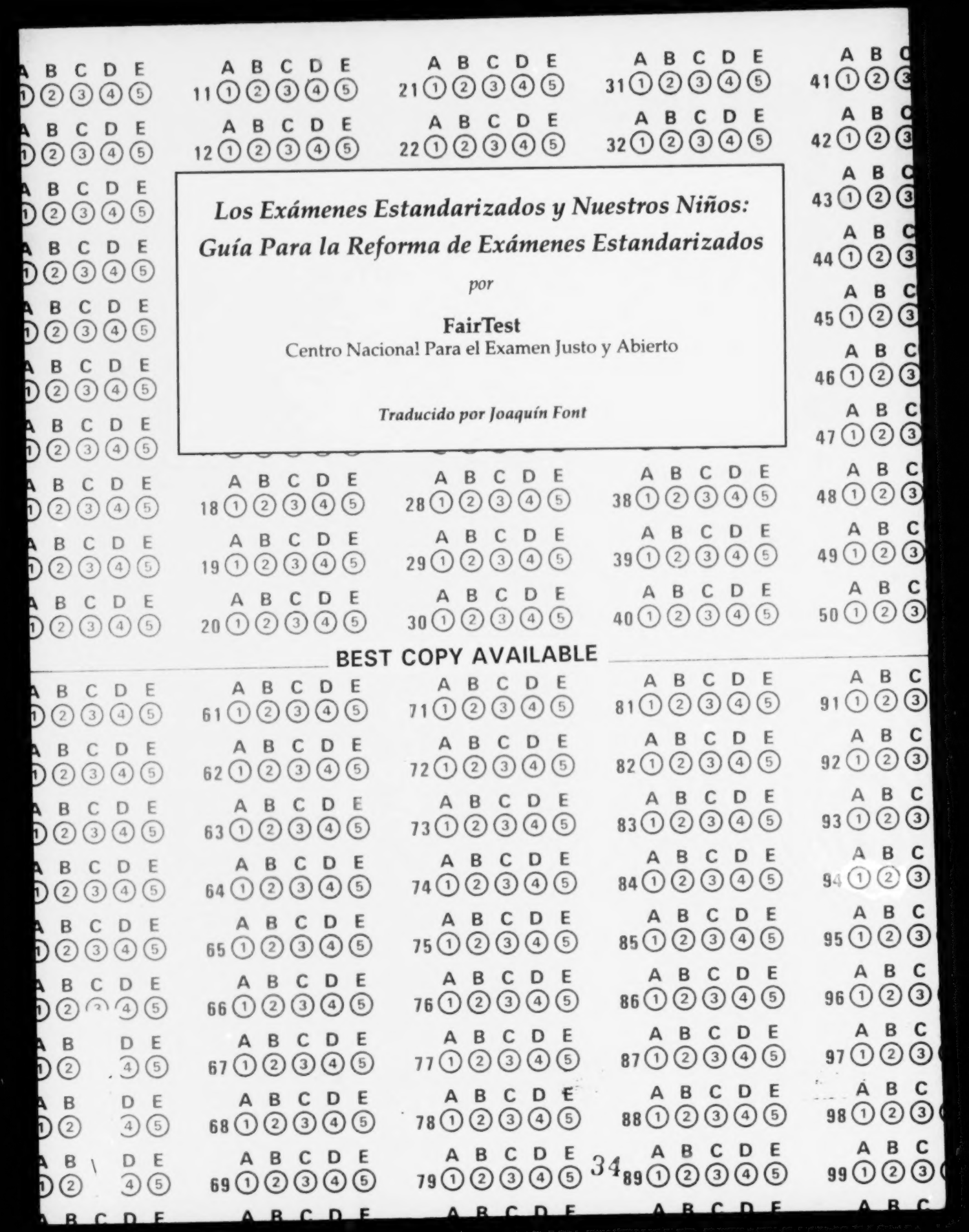
The National Center for Fair & Open Testing (FairTest) is an advocacy organization that works to end the flaws and misuses of standardized testing and to make certain that evaluation of students is fair, open, accurate, accountable, and educationally sound.

FairTest is grateful to the **Lilly Endowment** and the **Edward W. Hazen Foundation** for their generous support of this *Guide to Testing Reform.*

FairTest also thanks our other funders, the **Aaron Diamond Foundation**, the **New York Community Trust**, the **Golden Rule Insurance Company**, the **Ford Foundation**, the **Rockefeller Family Fund**, and the **Philip M. Stern Family Fund**, for helping to make *Standardized Tests and Our Children* possible.

Stamp

BEST COPY AVAILABLE



Los Exámenes Estandarizados y Nuestros Niños:
Guía Para la Reforma de Exámenes Estandarizados

por

FairTest
Centro Nacional Para el Examen Justo y Abierto

Traducido por Joaquín Font

BEST COPY AVAILABLE

FairTest

National Center for Fair & Open Testing

Primavera del 1991

Estimados Lectores:

Los exámenes estandarizados juegan un papel importante en la educación de nuestros hijos. Las puntuaciones obtenidas en estos exámenes pueden determinar a qué escuelas ellos asistirán, qué clases especiales deberán tomar, y si ellos serán seleccionados para cursos vocacionales o para los que preparan a quienes van a la universidad. Los resultados de los exámenes también pueden ser usados para ponerles a niños etiquetas tales como "superdotado", "impedido en el aprendizaje", o "retardado mental educable".

El uso extenso de los exámenes estandarizados comenzó en los años sesenta. Durante los años ochenta su uso aumentó aún más debido a que muchas personas pensaban que ellos eran una buena manera de medir la calidad de las escuelas y cuán bien estaban aprendiendo nuestros niños.

Más recientemente, padres, maestros, y expertos en la educación han llegado a comprobar como estos exámenes hacen más daño que bien. Demasiadas veces, los exámenes estandarizados de selección múltiple y de respuestas cortas proveen información falsa acerca de nuestros niños. Muchos de estos exámenes le hacen una injusticia tanto a los estudiantes que vienen de familias de minorías, como a los de escasos recursos económicos, e igualmente a aquellos que no hablan el inglés como idioma principal. Algunos exámenes también están prejuiciados en cuanto a sexo.

Más aún, el uso excesivo de los exámenes estandarizados le ha hecho mal a la calidad de la educación. Los maestros se han visto obligados a enseñar las destrezas limitadas que estos exámenes miden — y enseñarlas de forma limitada. Los maestros acaban por ignorar las "destrezas superiores de pensamiento" que los niños necesitan aprender en realidad.

De hecho, la administración de exámenes se ha convertido en un gran obstáculo para mejoras que son necesarias en las escuelas. La proliferación de exámenes también le ha restado el control de nuestras escuelas a los residentes locales y le ha trasladado ese poder a las grandes compañias examinadoras. Como resultado, los padres no tienen más sino menos información de utilidad sobre la educación de sus hijos.

A través del país hay más de cien millones de exámenes estandarizados que se dan cada año en las escuelas públicas. Eso significa que un estudiante típico debe tomar alrededor de tres docenas de exámenes antes de acabar la escuela superior. A muchos niños se les exige tomar su primer exámen cuando tienen tan sólo tres o cuatro años de edad.

Ahora, un creciente número de padres, maestros, administradores educacionales, líderes de derechos civiles y otros partidiarios han comenzado a trabjar en conjunto para cambiar esta situación. Se han formado grupos en muchos lugares para detener el uso erróneo y excesivo de exámenes fallidos. Salones de clase, escuelas, distritos, y estados a través de la nación están haciendo cambios en cómo evaluamos a los estudiantes. Existen métodos evaluadores que pueden ayudarle a los maestros a ser mejores evaluadores y proveer mejor información sobre el progreso estudiantil y la calidad de nuestras escuelas.

Apenas ha comenzado el movimiento por mejorar la evaluación — y con ello a las escuelas de la nación. Esta guía le ayudará comprender a los exámenes estandarizados y averiguar en qué se equivocan. De igual importancia, le presentará ideas sobre lo que usted puede hacer para cambiar la forma en que juzgamos a nuestros niños y a nuestras escuelas.

Esperamos poder trabajar con ustedes para hacer que nuestras escuelas sean mejores para todas nuestras escuelas.

Centro Nacional para la Evaluación Abierta & Justa (FairTest)

342 Broadway, Cambridge, Mass. 02139 (617) 864-4810 FAX (617) 497-2224

TABLA DE CONTENIDOS

I. Introducción 2

II. ¿Cómo Son Utilizados los Exámenes Estandarizados? 3

Breve Historia de la Administración de Exámenes

Algunos Exámenes Estandarizados Dados en la Nación

III. ¿Qué son los Exámenes Estandarizados? ¿Cómo Funcionan? 8

IV. ¿Qué Está Mal con los Exámenes Estandarizados? 11

¿Qué Derechos Tienen los Padres Sobre El Examen de Sus Niños?

V. Mejores Maneras de Evaluar a los Estudiantes 19

VI. Lo Que Usted Puede Hacer 22

Apéndice: Recursos que Usted Puede Usar 25

Apéndice: Significado Real de los Términos Examinadores (Glosario) 26

Declaración de la Campaña Nacional de FairTest Pro-Genuina Responsabilidad en la Educación 27

Aurora Gallagher (Austin, TX) y Kirk Anderson de *Cambridge Translation Resources* asistieron en la revisión de la traducción.

"Entre los exámenes de predisposición para la lectura, y las pruebas de aptitudes básicas, y los exámenes de I.Q. y los exámenes en sequencia sobre el progreso educativo y las pruebas de madurez mental y los exámenes de competencia mínima y las pruebas de promoción de grado y mi curso preparatorio para el PSAT, no hemos estado aprendiendo nadita en la escuela."

BEST COPY AVAILABLE

I. INTRODUCCIÓN

✎ *Miguelito apenas tiene cinco años. Esta semana él tomó su primera prueba estandarizada, un examen para decidir si él esta "Dispuesto" para el jardín de infantes. El examen dijo que él "No Está Dispuesto". Ahora Miguelito y sus padres y maestros todos creen que él es un fracaso. Pero en realidad es un niñito normal que fue puesto nervioso por un examen.*

✎ *Tamika tiene siete años. Acaba de obtener una puntuación en percentila de menos de treinta en un examen de lectura que es estandarizado y de selección múltiple. Su maestra dice que ella está progresando bien en la escuela, pero sus resultados en los exámenes indican que ella debe repetir el segundo grado. Si ella repite el año escolar habrá de hacer de nuevo todo el mismo trabajo. Tamika dice que se aburre en la escuela.*

✎ *Kim tiene diez años. Ella tuvo que repetir el tercer grado y fue puesta en una clase para estudiantes de "aprendizaje lento" debido a sus puntuaciones en los exámenes. Ahora es abril, casi la época de los exámenes de aptitud. Ella ha estado practicando destrezas para tomar exámenes desde septiembre. La maestra de Kim dice que ella no presta atención. Kim dice que ella odia la escuela porque "Solamente hacemos unas estúpidas hojas de trabajo. La escuela me hace sentir estúpida."*

✎ *Jaimito tiene quince años. Todos los miembros de su comunidad están asombrados ante su creatividad artística y su manera original e interesante de ver al mundo. Pero un examen de IQ dijo que él está "por debajo de lo normal". Así que fue colocado en un grupo de "aprendizaje lento", no obtuvo nunca buenas notas, y tuvo que repetir un año. Jimmy planea irse de la escuela tan pronto cumpla los dieciséis años.*

✎ *Juanita tiene diecisiete años. Ella siempre ha obtenido buenas notas. Esta semana recibió los resultados del Examen de Aptitud Escolástica (SAT) que ella necesita para solicitar a la mayoría de las universidades. Puntuó altamente en las matemáticas pero bajamente en lo verbal. Puesto que su primer idioma es el español, Juanita ha tenido que trabajar de forma especialmente dura para escribir bien el inglés. Aunque tenía que dedicarle bastante tiempo a este esfuerzo, ella mostró buenos resultados. Pero el SAT no le permitió el tiempo que necesitaba para pensar que palabra quería seleccionar o para corregir cuidadosamente a su trabajo. Sus puntuaciones verbales fueron tan bajas que quizás ella no conseguirá entrar a una buena universidad, y tiene poca oportunidad de ganar una beca.*

La administración de exámenes le ha hecho daño a todos estos niños. Algunos han sido colocados o "seleccionados" para clases aprisionantes que garantizan que ellos nunca llegarán a la universidad. Para demasiados, la escuela se ha convertido en un lugar frustrante y aburrido. Aún aquellos que, como Juanita, salen bien en la escuela, pueden averiguar a la larga que los exámenes le han hecho daño a sus oportunidades para tener éxito.

Estos ejemplos no son insólitos. Los exámenes estandarizados, de selección múltiple, y de respuestas cortas le hacen daño a muchos niños. Este folleto le mostrará a usted qué es un examen estandarizado, lo que significan en realidad los resultados de exámenes, y cómo pueden ser usados erróneamente. También le explicará a usted cómo y por qué estos exámenes le hacen daño a la educación aún cuando son utilizados de acuerdo con las intenciones de sus diseñadores.

Afortunadamente, existen mejores maneras de averiguar cómo se están desempeñando nuestros niños en las escuelas. Este folleto explica estos nuevos métodos de evaluación y muestra cómo pueden mejorar a nuestras escuelas.

Muchos padres, maestros, administradores de escuelas, y miembros preocupados de la comunidad han descubierto que es destructivo un exceso de exámenes estandarizados y de selección múltiple. En muchos lugares, las escuelas, distritos, e incluso los estados han comenzado a detener o reducir a los exámenes de selección múltiple, y a implementar mejores maneras de evaluar al aprendizaje estudiantil.

Nosotros podemos hacer que las escuelas sean un lugar donde todos los niños aprendan y se sientan felices. Cambiar la administración de exámenes es una fase necesaria para el mejoramiento de la educación. Juntos, podemos lograrlo.

II. ¿CÓMO SE USAN LOS EXÁMENES ESTANDARIZADOS?

En años recientes ha aumentado enormemente la administración de los exámenes estandarizados. Gran parte de este incremento se debe a preocupaciones sensatas por parte de padres y otros, ya que las escuelas no están educando bien a muchos de nuestros hijos. Se supone que los exámenes estandarizados nos digan cuán bien están aprendiendo nuestros niños. La teoría es que los resultados de los exámenes pueden motivar a los maestros y administradores para que mejoren su trabajo.

Aunque cada año se dan más y más exámenes estandarizados, no hay prueba de que hayan mejorado nuestras escuelas. De hecho, programas federales tales como el de la Evaluación Nacional del Progreso Educativo *(National Assesment of Educational Progress)* han demostrado que no ha habido casi ninguna mejora en el aprendizaje estudiantil desde el 1974, a pesar de que ahora se administran muchos más exámenes.

Los exámenes estandarizados son utilizados para un número de propósitos cada vez mayor. Algunos de los usos más comunes son:

Exámenes de Predisposición

También se conocen como "pruebas de madurez para el aprendizaje" o de "aprestamiento escolar". Estos exámenes son usados para decidir si un niño esta "dispuesto" para entrar al jardín de infantes o al primer grado. A menudo los niños evaluados como "no dispuestos" son excluídos de la escuela, obligados a repetir un año, o colocados en un jardín de la infancia o primer grado de tipo "transicional".

Desafortunadamente, la investigación demuestra que las puntuaciones en los exámenes de predisposición se equivocan incluso la mitad de las veces. Parte del problema está en pruebas de baja calidad que usan malas preguntas. Los niños tambien pueden desarrollarse en fases repentinas, y no siempre aprenden las cosas en un mismo orden. Un niño que es etiquetado como "indispuesto" hoy podría estar "dispuesto" el próximo mes. Los jóvenes y niños de familias de escasos recursos económicos son los que más probabilidad tendrían de ser etiquetados como "no dispuestos".

Hay evidencia que demuestra que son muy pocos los estudiantes que se benefician de ser excluidos de sus escuelas o mantenidos por un año adicional en un "programa transicional". Sencillamente, no es cierto que haya un gran número de estudiantes "no dispuestos" para la escuela. El verdadero problema está en las escuelas que están obligando a los niños a hacer tipos de trabajo equivocados y que no corresponden a sus edades -en parte con la intención de que ellos puedan puntuar mejor en los exámenes estandarizados que tomarán en grados superiores.

HI Y LOIS
¿SE ENCUENTRA MAL JASON?
YO CREO QUE SUFRE TENSIÓN.
BLÍ BLÍ BLÍ
MAÑANA ÉL TIENE UNA GRAN PRUEBA ESCOLAR.
POBRE NIÑO...
GULGO GLGLGLO
... SIN LA TETERA, NO PODRÁ CONCENTRARSE...
GLLLGA GLLLGA GLLLGA

Traducción libre. Reimpreso con el permiso especial de *King Features Syndicate, Inc.*

BEST COPY AVAILABLE

O sea que los exámenes de predisposición son usados para ver qué niños pueden cumplir con estas expectativas, las cuales son poco razonables.

Si las escuelas se "dispusiesen" para todos nuestros niños — en vez de los niños tener que estar "dispuestos" para la escuela — entonces ya no serían necesarios los exámenes de predisposición.

Pruebas de Discernimiento del Desarrollo

Se supone que los administradores de las escuelas le administren "pruebas de discernimiento del desarrollo" a niños menores únicamente para averiguar si deben ser referidos a un especialista en caso de que necesiten ayuda adicional. Los investigadores han encontrado que a menudo se equivocan las puntuaciones en estos exámenes. Con frecuencia estos exámenes son utilizados erróneamente sin que se obtenga la debida opinión de padres o maestros — de manera que se usan para colocar a los niños en programas, y no simplemente para que sean referidos para más estudio. La utilización de estos exámenes también fomenta el que los maestros y los trabajadores de salud mental se enfoquen tan sólo en las debilidades de los niños y que ignoren sus puntos fuertes.

Es cierto que los maestros necesitan saber cómo buscar señales de que un niño necesita alguna ayuda especial, pero para ello existen mejores métodos que los de la mayoría de los exámenes de discernimiento.

Exámenes de IQ

Las escuelas también utilizan exámenes del "cociente intelectual" (CI, o *IQ* que abrieva a *"intellgence quotient"*) de "habilidades cognoscitivas", y de "aptitud" para clasificar y etiquetar a niños cuando comienzan la escuela y, a menudo, también cuando entran a la escuela superior. Los exámenes de IQ asumen que la inteligencia es una sola cosa, en vez de comprender diversas habilidades. También asumen que la inteligencia es

Algunos de los Exámenes Estandarizados Dados en la Nación

Cientos de exámenes estandarizados diferentes son administrados cada año en las escuelas de la nación. Raras veces se les ofrece a los padres explicación alguna sobre el propósito del examen. En muchos casos los maestros, principales, y otros oficiales educacionales ni tan siquiera saben qué es lo que deben significar las puntuaciones de los exámenes. Aquí abajo citamos a tán sólo unos cuantos de los exámenes utilizados con mucha frequencia. Las abreviaciones se refieren a los nombres en inglés.

Exámenes de Predisposición
- Examen de Predisposición de la Escuela Gesell, Examen Pre-Escolar Gesell
- Examen de Predisposición Metropolitano
- Brigance (varios exámenes de predisposición escolar)

Exámenes de Cociente Intelectual, de Habilidad y de Aptitud
- Escala Wechsler Para Niños -Revisada (WISC-R)
- Escala de Inteligencia Stanford-Binet
- Examen Otis-Lenon para la Habilidad Escolar (OLSAT)
- Examen de Habilidades Cognoscitivas (COGAT)
- Examen de Destrezas Cognoscitivas (TCS)
- Serie Asesoradora de Niños Kaufman (K-ABC)
- Serie Psico-Educacional Woodcock-Johnson
- Exámenes de Aptitudes Diferenciales

Exámenes de Discernimiento
- Indicadores de Desarrollo para el Asesoramiento del Aprendizaje – Revisados (DIAL-R)
- Examen Denver Pro-Discernimiento del Desarrollo (DDST)
- Examen de Discernimiento McCarthy
- Instrumento de Mineápolis Pro-Discernimiento Escolar (MPSI)
- Exámenes de Logros K-12

Exámenes de Logros de California
- Exámenes de Destrezas Básicas de Iowa
- Examen Global de Aptitudes Basicas (CTBS)
- Examen Metropolitano de Logros (MAT o MET)

BEST COPY AVAILABLE

fija y permanente. Muchos investigadores no están de acuerdo.

Estudios demuestran que las puntuaciones de IQ pueden modificarse si hay un adiestramiento previo a la prueba, o se come un buen desayuno, o simplemente al tener personas más amistosas como administradores del exámen. Esto sugiere que los exámenes de IQ son meramente otro tipo de prueba de logros. Por lo menos dos cortes federales han ordenado a escuelas que dejen de utilizar los exámenes de IQ para colocar a niños en programas educacionales, porque estos exámenes están prejuiciados en cuanto a cultura y a raza.

Ya que no existe evidencia de que los exámenes de IQ puedan predecir cuan exitosa será una persona más tarde en su vida, no hay ninguna razón para requerir que todos los niños tomen a estos exámenes.

Nivelación en Trayectorias

Un sólo exámen puede determinar a la totalidad de la carrera educacional de un niño. Con frecuencia el personal educativo utiliza a los exámenes para "nivelar" a los estudiantes. Se les pone dentro de diferentes "trayectorias" de instrucción, designadas a veces como "grupos de habilidad". Según las puntuaciones, los estudiantes pueden ser consignados a programas de superdotados, accelerados, normales, remediativos, o de educación especial.

Supuestamente, la nivelación debe ayudar a los estudiantes al colocarlos en grupos que no son ni demasiado fáciles ni demasiado difíciles para ellos. Sin embargo, los investigadores educacionales han concluido que este tipo de colocación casi siempre le hace daño a los estudiantes de los niveles bajos. Y usualmente no les ayuda a los niños que son ubicados en las trayectorias altas a obtener mejores resultados que los que hubieran logrado de haber sido colocados en grupos de niveles mixtos.

En un sistema de nivelación los estudiantes en niveles más altos son los que tienen más probabilidad de trabajar en proyectos, leer libros, y discutir problemas. Debido a que su educación es más interesante y significativa, los de trayectorias más altas aprenden más y con mayor rapidez.

Las escuelas a menudo tienen bajas expectativas de aquellos estudiantes que han obtenido puntuaciones bajas en los exámenes estandarizados. Es probable que estos estudiantes reciban una educación inferior. Subir las puntuaciones de los exámenes pasa a ser el foco principal de su educación. Como resultado, gran parte del trabajo de los estudiantes de niveles inferiores consiste en memorizar y repetir la información que los maestros creen que necesitan los estudiantes con tal de obtener mejores puntuaciones en los exámenes.

Los estudiantes de trayectorias inferiores casi siempre se van quedando más y más retrasados. A menudo no acaban sus estudios.

BEST COPY AVAILABLE

La nivelación también puede tener como resultado la segregación racial de una escuela. A los afroamericanos, latinos e inmigrantes recientes se les obliga a entrar en los niveles más bajos tan sólo a causa de sus bajas puntuaciones. En muchos distritos escolares, los cursos avanzados están inundados de niños blancos que hablan inglés.

A los estudiantes con las puntuaciones más bajas se les puede poner de etiqueta algo así como "retardado mental educable" (*EMR*, para *"educable mentally retarded"*). Durante los años sesenta los afroamericanos constituían menos del 10% de la matrícula escolar total pero más del 25% de todos los estudiantes en los salones de clase para los "retardados mentales educables". Después de examinar toda la evidencia, una corte federal encontró que los exámenes estaban prejuiciados. El juez prohibió el uso de exámenes de cociente intelectual en California para "nivelar" a niños afroamericanos, colocándolos en las clases de retardados mentales educables.

Muchas escuelas y distritos están comenzando a cambiar sus normas con tal de reducir o detener a los nivelamientos. Si se detiene a la nivelación, entonces también puede y debe detenerse a la administración de exámenes que ha servido como justificación para esa nivelación.

Promoción de Grados

Algunos distritos ecolares requieren que un niño apruebe un examen estandarizado antes de que se le promueva al próximo grado. Los estudiantes que no aprueban el exámen son mantenidos en el mismo grado: son "retenidos" o "suspendidos". Se supone que esto es para que los niños tengan una segunda oportunidad para aprender el material.

Sin embargo, los estudios demuestran que la retención causa más problemas que los que resuelve. Suspender a un niño generalmente no conduce a un mejoramiento en el trabajo escolar. De hecho, un estudiante que es promovido en general logra mejores resultados que uno parecido a él pero que es retenido. Los niños que son retenidos a menudo sufren daños psicológicos y perdida de auto-estima. Los estudiantes que han sido suspendidos también tienen mucha más probabilidad de abandonar la escuela. Afortunadamente, ahora hay más escuelas que proveen ayuda adicional a estudiantes sin tener que retenerles.

A pesar de todo lo que se dice acerca del uso de pruebas para mejorar la calidad educacional, el principal propósito de los exámenes estandarizados sigue siendo el de clasificar a los niños, dividiéndolos en grupos. El uso excesivo de las pruebas obliga a las escuelas a utilizar prácticas — la clasificación, la suspensión, y la nivelación — que hacen más daño que bien.

"Tras 20 años escolares, su examen de aptitud demuestra que usted tiene habilidad en una sola cosa - tomar exámenes."

BEST COPY AVAILABLE

Breve Historia de los Exámenes Estandrizados

Al principio, el principal uso de los exámenes estandarizados era el de clasificar personas dentro de categorías. El primero de ellos fue el exámen de IQ (abreviación en inglés de "Cociente Intelectual", o CI en español), desarrollado hace cien años en Francia. Fue diseñado para descrubrir niños escolares que podrían necesitar ayuda especial y también fue usado para asignar a programas especiales a ciertos niños inmigrados de África y Asia. Así se garantizaba que ciertas escuelas permaneciesen con niños blancos.

A comienzos de siglo, el uso de los exámenes de IQ fue promovido en los Estados Unidos por los seguidores de la eugenesia -gente que creía que la inteligencia es heredada y que ciertos grupos (los del noroeste de Europa) eran, por su naturaleza, más inteligentes que otros.

Durante la primera guerra mundial los partidiarios de la eugenesia convencieron al gobierno para que experimentase con el uso de los exámenes de IQ para asignarles trabajos a los soldados. Estos exámenes eran extraños y sus resultados igual de raros. Miles de personas fueron etiquetadas como "morones" o aún peores clasificaciones simplemente porque no podían contestar preguntas tales como:

- **El número de piernas que tiene un kafir es: 2, 4, 6, 8**
- **Christy Mathewson tiene fama como: escritor, artista, jugador de beisból, comediante**
- **Karo es una: medicina patentizada, pasta dental, un desinfectante, o producto alimenticio**
- **Un disco de caucho es usado en el: tenis, futból norteamericano, hockey, golf**

Los exámenes no tenían relación con la capacidad para ejecutar ciertas tareas concretas. Algunos historiadores sostienen que estas pruebas fueron usadas raras veces para asignarles trabajos a los soldados.

Los examinadores fueron aún más lejos. Abogaron agresivamente por el uso de exámenes de este tipo para determinar qué emigrantes deberían ser admitidos a los Estados Unidos. Una vez más, grupos regionales y étnicos — tales como los provenentes del sur y este de Europa, africanos, asiáticos, latinos, y judíos — fueron designados como "flojos de mente" e inferiores. Las leyes de inmigración fueron cambiadas para mantener a esta gente fuera de los Estados Unidos.

Los productores de exámenes continuaron urgiendo un uso cáda vez más amplio de sus productos. En los años veinte se desarrollaron las pruebas de selección múltiple, basadas sobre las mismas ideas que las de IQ. En esa década (del 1920), los exámenes de IQ y de logros comenzaron a ser utilizados comúnmente para la selección y ubicamiento de niños en diversos grupos educacionales. Aunque muchas escuelas admninistraban ambos tipos de pruebas, el estudiante promedio tomaba tan sólo unos cuantos exámenes estandarizados a lo largo de sus doce años escolares.

En los años sesenta, el desarrollo de computadoras que podían calificar a miles de exámenes estandarizados por máquina les hizo más baratos de corregir. Casi a la misma vez, los oficiales públicos buscaban un método fácil y rápido para mejorar la educación y ganar así la carrera espacial con Rusia. Durante este tiempo, fue creciendo la influencia de las compañías examinadoras sobre la educacón. Sus vendedores convencieron a muchos distritos escolares para que adoptasen a los exámenes de logros de selección múltiple como la mejor forma de evaluar el éxito de sus estudiantes. También persuadieron al gobierno federal de que estos exámenes eran la mejor forma de medir el progreso de estudiantes en programas especiales tales como el "*Chapter 1*" (el "Capítulo 1" es un programa con fondos federales que le provee ayuda adicional a estudiantes de escasos recursos económicos que se retrasan en lo académico).

En los años ochenta, varios informes importantes concluyeron que muchos niños no estaban aprendiendo mucho en las escuelas. Una vez más, el incremento en la preocupación por lo académico condujo a un aumento en los exámenes estandarizados. En esta ocasión, las puntuaciones de los exámenes fueron utilizadas para evaluar programas, escuelas, y distritos enteros. Muchos estados comenzaron a requerir que sus estudiantes pasasen pruebas de "competencia mínima" para la promoción y la graduación. Incluso se les obligó a los maestros a tomar exámenes estandarizados.

Hoy en día se administran anualmente más de cien millones de exámenes en aulas por toda la nación. El estudiante promedio ha de tomar entre dos y cuatro pruebas cada año.

BEST COPY AVAILABLE

III. ¿QUÉ SON LOS EXÁMENES ESTANDARIZADOS?

A los exámenes se les llama "estandarizados" cuando le administran preguntas idénticas bajo condiciones similares a grupos grandes de estudiantes. La mayoría de los exámenes estandarizados son de selección múltiple y se corrigen con una computadora que ha sido programada para valorar a tan sólo una respuesta como la "correcta" para cada pregunta.

Usted probablemente tomó pruebas como estas cuando estaba en la escuela. Eran los exámenes que le hacían llenar blancos o oscurecer circulitos con un lápiz "Número 2".

Ciertos estados y algunos distritos escolares grandes desarrollan sus propios exámenes estandarizados, pero la mayoría son hechos por compañías privadas. Algunos manufactoradores de exámenes son empresas pequeñas. Sin embargo, muchos son negocios grandes. De hecho, el mercado de las pruebas administradas en las escuelas públicas está dominado por tres compañías enormes — la MacMillan/McGraw-Hill, la Harcourt Brace Jovanovich, y la Houghton Mifflin — que ganan decenas de millones de dólares anualmente trás la venta de sus exámenes. La mayoría de las pruebas de entrada a la universidad son hechas por el *Educational Testing Service* (Servicio Educacional de Exámenes), el cual tuvo ingresos de más de un cuarto de billón de dólares el año pasado.

Para que un examen sea preciso y significativo no basta con que se le llame "estandarizado". Las compañías examinadoras sostienen que sus pruebas estan diseñadas científicamente para medir la habilidad o logros de un niño. Pero la mayoría miden sólo lo bien que un estudiante puede seleccionar entre cuatro o cinco respuestas — y no sólo en términos de la rapidez con que responde sino también de acuerdo a la lógica limitada de la prueba.

Algunas veces se les llama "exámenes objetivos" a los que son estandarizados y de selección múltiple. Esto se debe a que el examen es corregido por una máquina o por hojas de puntuaciones — ningún individuo decide la puntuación del niño. Por otra parte, hay muchas otras razones por las cuales estas pruebas no son objetivas. Siguen habiendo seres humanos que escogen qué preguntas deben hacerse y con qué palabras han de ser expresadas. Las decisiones en torno a cuál es la contestación correcta, cuáles las incorrectas, y sobre cuál es la puntuación mínima para poder pasar el exámen — todo esto todavía está a cargo de los examinadores, no de máquinas. Todas estas decisiones claves son tomadas por individuos que son igual de subjetivos que los de cualquier otro grupo.

Dos Tipos de Examen Estandarizado

Las pruebas estandarizadas vienen en dos variedades principales: de referencia normativa y de referencia por criterio. Es importante saber un poco acerca de cómo funciona cada tipo de examen ya que sus resultados tienen significados muy diferentes.

Exámenes de Referencia Normativa (NRT)

Los exámenes de referencia normativa (*norm-referenced tests* o "NRT") comparan las puntuaciones de cada niño frente a un grupo de estudiantes que ya ha tomado el mismo examen. Al compararlo con éste otro grupo de estudiantes, la puntuación indica si el niño iguala, supera, o queda por debajo del promedio de esos estudiantes.

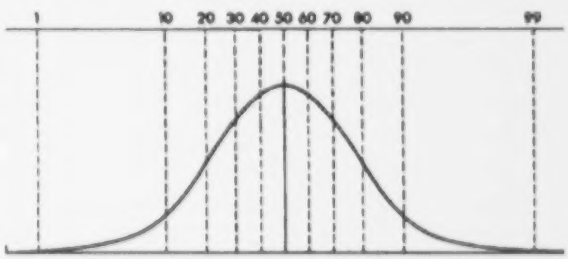
Los diseñadores de pruebas establecen un NRT de manera que la mitad de los estudiantes que lo toman deben terminar por debajo del promedio de las puntuaciones, y la otra mitad por encima. El esperar que todos los estudiantes estén por encima del promedio — o sea, sobre 50 en percentilas — es como si uno esperase que todos los equipos en una liga de béisbol tengan que acabar por encima del promedio. Tanto en los exámenes de referencia normativa como en el béisbol, para cada "ganador", tiene que haber un perdedor.

Los creadores de pruebas también han determinado que la mayoría de los estudiantes puntuan

alrededor de la mitad del promedio y que poquísimos logran resultados muy altos o muy bajos en los exámenes.

Ellos asumen que los resultados deben acabar organizados de acuerdo con una curva "normal" o con "forma de campana" (véase el gráfico debajo).

Rangos de Percentilas en una "Curva Normal"



Los números que atraviesan la parte superior indican el rango posible. Por ejemplo, todos los estudiantes a la izquierda del 1O puntuaron menos de la décima percentila. La mayoría de los estudiantes puntuan en los rangos intermedios -donde es más alta la curva.

Sin embargo, el conocimiento y habilidades de individuos no necesariamente encajan con este modelo. La curva "normal" existe porque le conviene a los examinadores, y no porque refleja la realidad.

Para producir el tipo de curva deseada por los examinadores, la prueba tiene que ser "normalizada". A un grupo pequeño de estudiantes se les administra muchas preguntas. Los diseñadores escogen entonces qué preguntas ("materias") formarán parte en el examen. Ellos retienen únicamente a aquellas "materias" que producirán resultados que puedan crear una curva. Algunas materias confusas o engañosas son retenidas, pero la mayor parte de las preguntas que son acertadas por los estudiantes son rechazadas.

Al construir un NRT, tiene mayor importancia escoger preguntas que seleccionen a estudiantes que ocuparse de que sea adecuado el contenido cubierto por el examen.

La selección de puntuaciones conforme a la curva permite que los resultados de los exámenes sean comparados. Así pues, se dice que Sally puntuó en la percentila 75. Esto significa que una cuarta parte (25%) de los estudiantes de su mismo grado que tomaron el examen puntuaron igual o mejor que ella.

Exámenes Referidos a un Criterio **(CRT)**

Los exámenes referidos a un criterio (CRT) no comparan los estudiantes los unos con los otros. Simplemente miden si un niño ha amaestrado a un objetivo específico de aprendizaje. El examen de selección múltiple que toman muchas personas para obtener la licencia de conducir (de manejar) es un ejemplo de un examen referido a un criterio.

Un examen no es mejor simplemente porque se refiera a un criterio. Pero supongamos que un examen de manejar preguntase acerca de cómo funciona un motor de automóvil o que estuviese escrito en una lengua desconocida por el conductor.

Desafortunadamente, la mayoría de los CRT son esencialmente iguales que los exámenes referidos a una norma. Causan el mismo mal a los niños y educadores que hacen los que se refieren a normas. Las pruebas estandarizadas no tienen que ser de selección múltiple o de respuesta corta. Pueden permitir respuestas en forma de ensayo o pedirles a los estudiantes que realizen tareas o que hagan un producto. Por ejemplo, se les puede pedir a estudiantes de ciencias que ejecuten y expliquen un experimento de laboratorio, en vez de limitarlos a contestar con respuestas verbales. De hecho, el estado de Nueva York tiene un examen de laboratorio estandarizado que funciona de esta manera. (Véase la sección V, que cubre mejores maneras de evaluar a los estudiantes.)

LO QUE SIGNIFICAN LAS PUNTUACIONES DE LOS EXÁMENES

✎ La **Puntuación Inicial** es el número total de preguntas que un niño o niña saca bien en un examen.

✎ Un **Porcentaje Correcto** en la puntuación se crea al dividirse el número de respuestas acertadas por el total de las preguntas. Por ejemplo, un niño que acierte 10 respuestas de un total de veinte problemas tendría un porcentaje correcto de un 50%.

✎ Una **Puntuación Percentila** o **de Rango** le dice cómo compara la puntuación del niño con la de los estudiantes en base a los cuáles fue normalizada la prueba. Esto no significa lo mismo que el porcentaje de preguntas que él o ella saca bien. Una percentila o rango de 50 significa que la puntuación del estudiante fue mejor que la del 50% de los estudiantes en la muestra normalizadora (véase el gráfico de la página 9).

✎ Una **Puntuación Estandard** esta diseñada para mostrar cuánto difiere la puntuación de un estudiante de la del promedio. Tomando casos como el del examen de I.Q. — una prueba cuya puntuación promedio es de 100 — un resultado de 110 supera al promedio por 10 puntos mientras que uno de 90 estaría por debajo del promedio por esa misma cantidad.

✎ La expresión **"Puntuacion Estanina"** proviene de la palabra *stanine* en inglés, la cuál es una abreviación para ***"standard nine"*** (estandarte nueve). Las estaninas dividen los resultados de un examen en nueve grupos, asignándoles el 1 al más bajo y el 9 al más alto. Se consideran como promedio a las puntuaciones 4, 5, y 6.

✎ Una **Equivalencia de Curva Normal (NCE)** es otro tipo de puntuación estandarizada, con un promedio de 50 y un márgen de puntuaciones que se extiende del 1 al 99. En cualquiera de estos exámenes, cada grado tiene su propio NCE. Si un niño tuvo un NCE de 45 en el grado segundo y de 55 en el tercero, ha progresado según el examen ya que sus resultados pasaron de estar por debajo del promedio a estar por encima del mismo. Las puntuaciones del *Chapter One* (Capítulo Uno) se consiguen con este sistema.

✎ Una ***Puntuación de Equivalencia de Grado*** intenta comparar a los resultados de su niño con los del mes y grado en los cuales estos equivaldrían a los del promedio. Según este modelo, un tomador de examen típico que rinde bien en el mes 5 del segundo grado obtendría una puntuación de 2.5. Utilizando este método, la mitad de los estudiantes que toman el examen a nivel nacional deben terminar por encima del promedio de su grado, y la otra mitad por debajo. Las puntuaciones de equivalencias de grado son engañosas con respecto a los estudiantes que salen o muy bien o muy mal en el exámen. Por ejemplo, un estudiante en el quinto mes del segundo grado que puntúa un 4.9 no está haciendo trabajo de cuarto grado, y es casi seguro que fracasaría si se le colocase en el cuarto grado. Él o ella tan sólo está obteniendo muy buenos resultados según las normas del examen.

Dadas las diferencias que pueden darse de un día para otro en la calidad del rendimiento de un niño en una prueba, y también el que cada examen tiene un **error de medición,** todas estas estadísticas deberían ser presentadas a los padres como **alcanzes de puntuación** o como **márgenes de confianza.** Estos cálculos muestran la probabilidad de que cambie la puntuación de su niño si volviese a tomar otra vez al mismo examen. De hecho, las puntuaciones pueden variar aún más de lo indicado por un márgen de confianza.

Todos los exámenes tienen un **error de medición**. Tomemos el caso de un estudiante que puntúa un 57 en un examen con referencia a una norma, con un alcanze de puntuación de 0-100 y una curva normal. En dos de cada tres casos su puntuación recaerá entre 50 y 64 — que sería su ***márgen de confianza*** o ***alcanze de puntuación.*** Se supone que el márgen de confianza le indique cuánto cambiaría la puntuación de su niño si él o ella volviese a tomar de nuevo al examen. Pero de hecho, las puntuaciones pueden variar aún más de lo que el márgen de confianza señala como probable. En este ejemplo, **una de cada tres veces, la puntuación estaría por debajo de 50 o por encima de 64.**

IV. ¿QUÉ ESTA MAL CON LOS EXÁMENES ESTANDARIZADOS?

Los defensores de los exámenes estandarizados afirman que el problema no radica en los exámenes, sino en cómo estos son usados. De seguro, las puntuaciones de estos exámenes con frecuencia se usan incorrectamente. Pero aun al disponerse de los mismos según lo requieren los examinadores, estos exámenes pueden causar un daño muy grande. Esto se debe a que los exámenes estandarizados están fundamentalmente defectuosos. En esta sección hemos de demostrarle el cómo y el por qué de este hecho.

Los resultados de los exámenes son a menudo incorrectos

Los examinadores sostienen que los exámenes estandarizados miden de forma útil y precisa los logros, habilidades o destrezas de un niño. En realidad, los exámenes estandarizados vigentes casi siempre no resultan ni *confiables* ni *válidos*.

Un examen sería completamente *confiable* si usted pudiera obtener iguales resultados al administrársele este examen a un mismo estudiante por segunda vez. Los resultados para una mayoría de estos exámenes varían ampliamente dependiendo de las condiciones de la prueba, de cómo se siente el estudiante ese día, de la manera en que son dadas las instrucciones o de cómo ha sido entrenado el estudiante. De ahí que la mayor parte de las puntuaciones obtenidas por individuos resulten erróneas (véase al respecto el "error de medición" en la página 10). Estos factores afectan especialmente a los más menores, por lo que sus exámenes resultan aún menos confiables.

Debido a estos problemas de confiabilidad (el

Preguntas de Examen Defectuosas

Las siguientes preguntas provienen de exámenes de logro por materia todavía vigentes. Estos artículos defectuosos demuestran parte del problema con el uso de preguntas de selección múltiple.

"La carne que estaba en _____ se quemó".

• el caliente • el horno • la cocina • la cena

(*Metropolitan Achievement Test* — Exámenes de logro por materia metropolitanos, sexta edición, segundo grado, 1985)

La respuesta deseada es horno. ¿Por qué es "la cena" incorrecta?

Identifique el dibujo de aquella persona que lo más probablemente provenga de otro país.

(*Comprehensive Tests of Basic Skill* — Exámenes Comprensivos de Destrezas Básicas, 4ta Edición, grados 1.0 - 2.2, 1989)

¿Por qué cualquiera de estos niños debería "provenir probablemente de otro país"?

BEST COPY AVAILABLE

¿Qúe tal su I.Q.?

"¿Cuál es la estatura de un americano promedio?"

Toda respuesta entre 5'7" y 5'11" es correcta

Los niños responden a menudo citando la estatura de sus padres. Los asiáticos y los latinoamericanos tienden a ser más bajos, mientras que otros hombres son más altos. ¿Cómo puede la respuesta de un niño a tal pregunta reflejar su "inteligencia" o la carencia de ésta?

¿Qué debe hacer uno de cortarse un dedo?

Una respuesta que amerita 2 puntos:
"Ponerse una curita en la herida".
Una respuesta que amerita 1 punto: "Ir al doctor (u hospital) ... para que le tomen puntos a la herida."
Una respuesta que no amerita ningún punto: "Chuparse la sangre ... No sobrecogerse de pánico"

Los niños minoritarios tienden a hacer muy mal en esta pregunta. Un sociólogo de Baltimore, Maryland, le preguntó a varios jóvenes de sectores urbanos pobres porqué respondían de la forma que lo hacían. Muchos de ellos respondían "ir al hospital" porque pensaban que "cortarse" significaba darse un tajo bien grande. Cuando a estos jóvenes se les dejaba saber que "cortarse" se refería a un tajo pequeño, casi todos respondían entonces, "ponerse una curita".

Ambas preguntas son del *Wechsler Intelligence Scales* para Niños — el examen de I.Q. Revisado (WISC-R) y todavía vigente en muchas escuelas neoyorquinas.

"error de medición" de un examen), casi todos los expertos concuerdan en que las puntuaciones de estos exámenes nunca deben constituir la base única o primaria para tomar una decisión acerca de un niño.

La *validez* de un examen es igualmente importante. Ésta se refiere a si un examen realmente mide lo que pretende medir. La información provista sobre la validez de los exámenes por parte de quienes los hacen es a menudo incompleta y engañosa. El que un examinador nos diga que un examen por él hecho ya ha sido validado no significa que su examen lo sea realmente ni que éste resulte lo suficientemente bueno como para usarlo en la toma de decisiones acerca de unos niños.

Un examen que se vende como válido para medir algo puede ser que mida una cosa diferente. Por ejemplo, un examen de "composición escrita" que le pide a un estudiante el marcar la palabra incorrecta en una oración no mide su capacidad para escribir: éste tan sólo indica cuán bueno es el estudiante en la corrección de pruebas. A esto se le añade que unos exámenes validados para cierto uso, como el de "discernimiento de desarrollo", pueden resultar inválidos si se emplean con un propósito diferente, por ejemplo en la "distribución de pupilos." Tales usos erróneos son muy comunes a través de la nacción.

Un examen supuestamente válido también puede ser que mida de forma deficiente la asignatura que pretende examinar. Por ejemplo, un examen de biología que en su mayor parte sólo le exige a los estudiantes el definir palabras ofrece una muestra muy pobre de aquello en lo que consiste realmente la materia de la biología.

Un examen que se supone pronostique de forma válida cuán bien hará un estudiante en el futuro sólo puede servir con precisión dependiendo de como las escuelas usan sus resultados. Por ejemplo, unos estudiantes con bajas puntuaciones en los exámenes podrían ser asignados a "una trayectoria de nivelación más lenta". El hecho de que estos niños no salgan bien en exámenes subsecuentes se debería a su pobre educación en el programa de nivelación y no a su falta de talento.

Debido a su confiabilidad y válidez inadecuadas, ninguna puntuación de examen resulta lo suficientemente precisa para ser usada como la base única o primaria al decidir sobre la colocación de un estudiante o su promoción de grado. Desafortunadamente, las escuelas a menudo toman estas decisiones en base a tales puntuaciones.

Muchos exámenes están prejuiciados

Todo examen refleja el trasfondo de quienes lo diseñan. La mayoría de los que preparan estos exámenes son profesionales blancos, de clase alta o media. Estos exámenes reflejan sus propias experiencias, no aquellas de la mayor parte de los niños que toman el examen.

Esto no significa que quienes fabrican estos exámenes son racistas, sexistas o elitistas. Ellos son simplemente humanos y limitados por el hecho de que ningún examen puede tomar en consideración todas las culturas diferentes que componen una sociedad tan diversa como la nuestra.

Todo artículo escrito de una forma que favorece a un grupo de estudiantes pondrá a otros en desventaja. Preguntas en tópicos de interés particular para los afroamericanos, los latinos o los asiáticos podrían ayudar a que éstos aumentasen sus puntuaciones respectivas, pero tales artículos son raramente incluídos. Por eso, el decir que unos niños tienen "poco talento" debido a que no tienen el trasfondo o experiencias de los niños blancos de clase media es claramente injusto.

Gran parte de estos prejuicios proviene del lenguaje usado en los exámenes. Las pruebas son escritas a menudo con una gramática muy compleja y palabras poco comunes. Este "inglés super-estándar" puede que ayude a incrementar las diferencias entre las puntuaciones, pero no es justo para la mayoría de los que toman estos exámenes.

Los exámenes también pueden estar prejuiciados sexualmente. Entre nuestros niños y niñas más jóvenes, algunos exámenes están prejuiciados en contra de los niños. Entre adolescentes, el prejuicio se inclina en la dirección opuesta; las jóvenes tienden a obtener puntuaciones más bajas en los exámenes de admisión universitaria estandarizados, a pesar de que ellas obtienen notas más altas que las de sus compañeros en escuela superior y en universidad.

Los que hacen estos exámenes dicen que los niños de bajos recursos y de grupos minoritarios obtienen puntuaciones más bajas debido a que su instrucción elemental con frecuencia no es tan buena. Esto es parcialmente cierto, pero el problema no se resuelve con unos exámenes prejuiciados, no confiables e inválidos que son los que determinan la asignación de nuestros niños a sus programas respectivos de nivelación. *De hecho, el uso excesivo de exámenes estandarizados a menudo empeora todavía más la situación de aquellos estudiantes que ya están en escuelas inferiores.*

Los exámenes a menudo son administrados inadecuadamente

El empleo de examinadores inexpertos, una asignación de asientos inadecuada y el uso de salones demasiado calientes, fríos o ruidosos pueden todos contribuir en la obtención de bajas puntuaciones. Su impacto varía entre diversos tipos de estudiante. Los niños más pequeños son los más afectados. Se ha demostrado que el uso de examinadores poco familiares con los exámenes afecta de forma negativa las puntuaciones de minorías o de jóvenes con escasos recursos económicos tomando estas pruebas. *Estos problemas hacen aún menos confiables los resultados de los exámenes, pero esta situación se ignora a menudo al considerar las puntuaciones de los niños.*

Los exámenes y las puntuaciones pueden estar equivocadas

Como todo ser humano, los que escriben los exámenes también cometen errores. A veces una pregunta puede tener más de una contestación "correcta". Otras veces, ninguna de las alternativas es correcta del todo. La respuesta correcta que tiene en mente el que escribe un examen puede que no sea siquiera la mejor contestación posible. A pesar de esto, un niño que marca dos alternativas o deja la pregunta en blanco pierde puntos de cualquier manera.

Las máquinas que computan las puntuaciones de unas hojas de respuesta también pueden cometer

¿Qué derechos tienen el padre y la madre sobre los exámenes de sus hijos?

Desafortunadamente, como padre usted tiene muy poco control sobre los exámenes de sus hijos. Bajo la ley estadounidense, a usted se le permite ver las puntuaciones de su hijo y demás expedientes que la escuela mantiene sobre el niño. Usted tiene derecho a prender una nota en toda información que reciba sobre un examen para protestar cualquier cosa que usted considere errónea.

Como padre, usted también tiene derecho a insistir que su hijo o hija sea examinado en su primer idioma si éste no es el inglés. Si su hijo tiene algún tipo de dificultad en el aprendizaje, usted puede tener derecho a insistir que se le provea ciertos acomodos especiales, tales como el que su exámen le sea leído en voz alta o el que se use una versión del mismo en letra más grande. Aparte de las anteriores, la ley le permite muy pocos derechos a los padres sobre los exámenes.

Los padres deberían tener muchos más derechos. Usted debería exigir que la escuela o el distrito de su hijo adoptaran los siguientes derechos de padre y estudiante si estos no han sido adoptados todavía:

✎ Como padre, usted *debería* tener derecho a ver un examen y revisar las preguntas, las respuestas deseadas y las respuestas de su hijo con un maestro o un administrador. Algunos estados han pasado una ley requiriéndole a los manufacturadores de los exámenes de admisión a la universidad que provean copias del tipo de preguntas que aparecen en tales exámenes con sus respuestas correctas para el uso de todo estudiante que así lo desee. Esta protección no aplica, sin embargo, a los exámenes administrados en escuela pública.

✎ Como padre, usted *debería* tener derecho a que se le expliquen los exámenes y los términos examinadores, de forma que usted pueda entenderlos. Las escuelas deberían informarle sobre cuándo ocurren los exámenes, qué tipos de pruebas se administran, y sobre cómo serán utilizados los resultados. También deben darle la puntuación de su niño y contestarle cualquier pregunta que usted pueda tener.

✎ Como padre, usted *debería* ser capaz de retar la puntuación de un examen si ésta no le parece correcta. Recuerde, usted también puede prenderle una nota a la hoja de puntuaciones explicando por qué es que ésta está mal.

Si usted se opone a una prueba, usted puede intentar que a su hijo se le exima de la misma. Desafortunadamente, si todos los demás niños están siendo examinados, usted no puede rehusarse legalmente a que su hija o hija sea examinado. Para todo examen especial, usted debe insistir que la escuela explique el examen y como sus resultados podrían ser usados por ella, antes de dar su consentimiento.

Como padre, usted tampoco tiene muchos derechos sobre como se usan los exámenes para situar a los niños en la escuela. La ley federal requiere que toda escuela obtenga su consentimiento para poder colocar a su hijo en cualquier otro programa que no sea el regular. Todo programa de seguimiento o nivelación es considerado parte del programa regular. La asignación de un estudiante a unas *clases de educación especial* o *de educación bilingüe* no lo es, sin embargo. Hubo un caso en el que el Departamento de Educación del Estado de New York dictaminó que los padres podían rechazar la decisión tomada por una escuela de asignar a su hijo a un programa de kindergarden o primer grado *transicionales*.

Si usted cree que su hijo ha sido asignado a una clase o a un programa de nivelación incorrecto debido a su puntuación en estos exámenes u otra razón cualquiera, usted debe retar la decisión. Debe investigar cuidadosamente toda colocación especial que la escuela proponga para su hijo. Estas decisiones *nunca* deberían ser tomadas primordialmente en base a las puntuaciones de un examen — ningún examen es lo suficientemente bueno. Para tomar decisiones educacionales sobre todo estudiante debería servir de base una información más amplia sobre como el mismo se desenvuelve en la escuela, en su casa y en su comunidad — una información obtenida a través de sus padres, de sus maestros y demás fuentes fidedignas.

Como padre, usted debería insistir en que se convocaran reuniones donde poder informarse acerca de los exámenes que su escuela emplea. En esas reuniones, los administradores podrían ponerle al día sobre sus prácticas de evaluación y usted podría cerciorarse por medio de preguntas de que los exámenes no están siendo usados incorrectamente. Usted también podría usar estos foros en colaboración con sus simpatizantes dentro del personal de la escuela para persuadir a otros padres y oficiales en la misma a que busquen maneras de cambiar sus prácticas de evaluación.

errores. Éstas pueden confundirse por una marca fuera de sitio en la página, alguna falla eléctrica u otro problema. El encontrar el circulito correcto para marcar una respuesta es particularme confuso para un niño pequeño. *Como consecuencia, las puntuaciones reportadas a un estudiante — o a una clase entera — pueden estar equivocadas.*

Los exámenes estandarizados penalizan la imaginación y el pensamiento creativo

Un niño que piensa creativa, cuidadosa o profundamente puede tener problemas con exámenes de selección multiple. Estos exámenes recompensan a los estudiantes que seleccionan rápidamente la "mejor" contestación sin pensar demasiado. Un niño que piensa por qué una alternativa es mejor que otra puede que no sea capaz de completar el examen o que se vea forzado a apurarse al final. Cualquiera que sea el caso, el resultado será una puntuación baja.

Algunos niños tienen dificultades para pensar de la manera que quieren los que prepararon el examen. Ellos sencillamente no son buenos en el juego de la selección multiple. Otros estudiantes se ponen nerviosos debido a la presión del tiempo. Algunos llegan a enfermarse de verdad, los más pequeños especialmente.

Las destrezas requeridas para la toma de un examen son las opuestas de los buenos hábitos de estudio que nuestras escuelas deberían enseñar. Cuando el objetivo de la enseñanza es obtener altas puntuaciones en los exámenes, la imaginación y el pensamiento creativo son desterrados del salón de clase.

Los exámenes no miden las destrezas más importantes del pensamiento

Los exámenes estandarizados de selección múltiple no bregan con el razonamiento, la capacidad de resolver problemas o la formación de ideas — "las altas destrezas del pensamiento". Estos se enfocan por el contrario en tareas simplistas como el reconocer datos. Éstas no son las destrezas requeridas para producir científicos, artistas, maestros o los demás adultos que nuestra sociedad necesita.

Los exámenes de selección múltiple asumen que todo problema sólo tiene una contestación correcta. La verdad es que muchos problemas tienen más de una solución. En el mundo real, el primer problema es con frecuencia poder determinar cual es el problema mismo. Los exámenes de selección múltiple le dicen cual es el problema sin permitirle que usted mismo lo determine. No hay materia en la que un examen de selección múltiple pueda evaluar adecuadamente las más altas destrezas del pensamiento o la capacidad para emplear el conocimiento en la solución de problemas reales. Un mecánico, por ejemplo, que no pasa un exámen de selección múltiple en mecánica aún así puede arreglarle su carro. Ningún examen estandarizado mide las abilidades de las personas, la imaginación, la creatividad, la determinación o el liderazgo.

Aun cuando puede ser que su niño no salga bien en los exámenes estandarizados, él o ella es capaz de salir bien en la escuela y más adelante en su vida. La mayoría de lo que nuestros niños necesitan saber no puede medirse con un examen estandarizado.

Los exámenes no reflejan cuánto es lo que la gente realmente aprende

Las ideas que subyacen los exámenes de logros por materia fueron desarrolladas por psicólogos del siglo diecinueve. Ellos creían que el aprendizaje era un proceso en el cual la persona absorbía pasivamente pequeños fragmentos de información individuales hasta compilar los suficientes fragmentos como para formar una idea completa.

Los psicólogos ahora saben que el aprendizaje es un proceso activo. Las personas — inclusive los niños pequeños — aprenden estableciendo conexiones entre lo que ya conocen y lo que intentan aprender. Si ellos no pueden participar activamente en la comprensión de lo que se les está siendo enseñado, esto les impide que aprendan o recuerden como debieran. Aun el aprendizaje de destrezas básicas — como hablar y escribir —

envuelve un modo de pensar complejo. Los estudiantes pueden y deben fortalecer sus destrezas básicas y las de más alto nivel a un mismo tiempo.

A pesar de que ha progresado enormemente nuestra comprensión del cerebro y de como uno aprende, los exámenes estandarizados siguen siendo los mismos. Hoy en día, la mayor parte de los educadores — aun los que hacen estas pruebas — concuerdan en que las escuelas no deben hacer que el currículo y la enseñanza se asemejen a exámenes. Cuando la enseñanza se encuentra controlada por exámenes, los estudiantes se vuelven pasivos e irreflexivos. Terminan por olvidar la mayor parte de lo que han "aprendido" pues esto último no tenía ningún significado para ellos. Los estudiantes no sólo dejan de adquirir destrezas de un más alto nivel: ellos ni siquiera aprenden las destrezas más básicas como debieran.

El uso de exámenes y métodos de evaluación que incitan al estudiante a pensar también pueden ayudar a que los maestros mejoren la calidad de su enseñanza. En la sección V, discutimos estas mejores maneras de evaluación.

Las puntuaciones de los exámenes pueden que no reflejen lo que un niño en realidad ha aprendido

Un examen estandarizado está diseñado para ser dado a niños de toda la nación. Su contenido puede que tenga poco que ver con lo que se enseña en una escuela cualquiera en particular. No todas las clases cubren su material en el mismo orden. Si los niños son examinados en un material que su maestro todavía no les ha enseñado, sus puntuaciones sufrirán por ello.

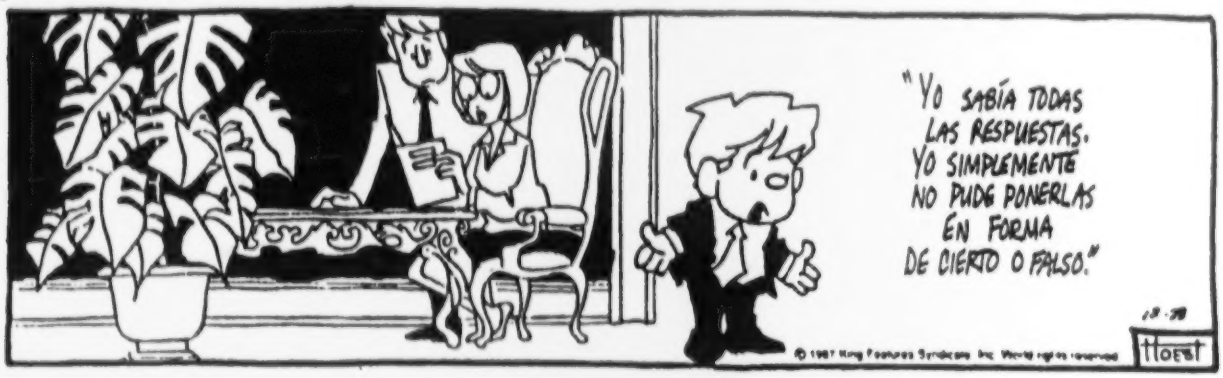
LOS EXÁMENES LE HACEN DAÑO A TODA LA EDUCACION

Los exámenes de selección múltiple se han apoderado de la educación. A pesar de las fallas de los exámenes y el daño que le causan a millones de niños, su uso ha llegado a controlar lo que las escuelas cubren y lo que los maestros enseñan. Un mayor número de padres, maestros y demás personas han llegado a ver, sin embargo, que esto constituye un error muy grave. El depender excesivamente en los exámenes no mejora la educación. De hecho, previene las mejoras requeridas y les confiere un poder a las compañías examinadoras sobre las cuales no tiene control padres ni maestros ni comunidades.

Los exámenes estandarizados perjudican la calidad de una escuela

Las puntuaciones de los estudiantes en sus exámenes son usadas cada vez más para evaluar a maestros, escuelas y distritos enteros. A través de la

¡Qué muchacho!/Bill Hoest



Reimpreso con el permiso especial de *King Features Syndicate, Inc.*

BEST COPY AVAILABLE

nación, los periódicos divulgan los promedios de puntuación como si éstos fueran una medida precisa de la calidad de una escuela. En ciertos estados, la cantidad de fondos asignados a una escuela depende de los resultados que sus estudiantes obtienen en estos exámenes.

Todos estos factores crean situaciones de exámenes de "alta transcendencia". Cuando los oficiales y los maestros de una escuela se sienten presionados a tener que mejorar estas puntuaciones, comienzan a "enseñar para el examen". Esto significa que los salones de clase se convierten en centros de entrenamiento para los exámenes.

Algunas escuelas comienzan en septiembre a preparar sus estudiantes para los exámenes que han de ser dados en abril y mayo. Muchas escuelas comienzan el entrenamiento en enero. Esto resulta a veces en puntuaciones más altas, pero toda esta preparación le roba tiempo a los niños del que se requiere para el estudio más serio.

El "enseñar para un examen" ignora casi todo lo que no se incluye en exámenes estandarizados de selección múltiple. Si un examen de lectura sólo hace preguntas basadas en párrafos cortos, esto será en lo que el maestro se enfocará. Las novelas, los ensayos y los poemas a menudo no son estudiados. Muchas veces, los estudiantes en tales clases no leen libros, no trabajan en proyectos, no resuelven problemas de la vida real ni llegan a comprender la ciencia o la historia. En vez de esto, sólo se dedican a memorizarse datos aislados y hacer ejercicios en hojas de trabajo que parecen exámenes. Esto ocurre mayormente en los programas de nivelación donde es más probable que los estudiantes provengan de familias con pocos recursos o de grupos minoritarios. Irónicamente, estos niños tienden a estar en programas de nivelación debido a las mismas puntuaciones de sus exámenes.

Los libros de texto también tienden a ser "simplificados" para que su forma y contexto correspondan al de los exámenes. A menudo, las mismas compañias que hacen estos últimos, también escriben los libros de texto, diseñandolos de forma tal que puedan ayudar a subir las puntuaciones de los exámenes. Algunos textos llegan a incluir un examen de selección múltiple al final de cada capítulo. Los estudiantes se vuelven, como resultado, menos interesados en la lectura y apenas recuerdan lo que han leido. Aburridos y frustrados, son estos los estudiantes que lo más probable se darán de baja de la escuela.

Los partidarios de estos exámenes alegan que el enfocarse en los mismos conduce al mejoramiento de destrezas básicas. Ellos señalan las puntuaciones crecientes en los exámenes de logro por materias como evidencia. Desafortunadamente, unas puntuaciones crecientes no indican que se haya aprendido más o que la calidad de la educación haya mejorado. El entrenamiento para un examen puede aumentar las puntuaciones en el mismo sin que se adquiera más conocimiento substancioso.

A pesar del gran aumento en el uso de exámenes estandarizados de selección múltiple durante los ochenta, los estudios del *National Assessment of Educational Progress* (Evaluación Nacional del Progreso Educativo) demuestran que la tasa de logros académicos ha aumentado apenas o quizás nada. De hecho, los niños en el cuarto grado en 1988 — los mismos estudiantes que se vieron más afectados bajo un currículo dominado por exámenes — leían mucho más mal que los del cuarto grado en 1980. Toda la evidencia disponible demuestra que los estudiantes no están desarrollando sus destrezas de más alto nivel en la manera que deberían. *Cuando los exámenes controlan las escuelas, quizás ascienden sus puntuaciones, pero la calidad de la educación usualmente empeora.*

Los exámenes no nos ofrecen ningún criterio de confiabilidad

Aquellos que favorecen los exámenes dicen que estos son necesarios por su confiabilidad. ¿De qué otra forma, ellos se preguntan, podemos saber si nuestros niños están aprendiendo? La confiabilidad

de un examen constituye una preocupación importante y legítima de padres, ciudadanos y legisladores. *El emplear una herramienta tan dañina para la educación y que yerra a menudo, no es forma, sin embargo, de satisfacer este criterio de confiabilidad genuina.*

Algunas escuelas hacen trampa debido a la presión de tener que aumentar las puntuaciones de los exámenes. Algunos oficiales escolares han sido cogidos distribuyendo respuestas a los estudiantes, enmendándoles sus errores y deshaciéndose de las hojas de respuestas más pobres. Se ha llegado a sugerirles a los estudiantes con un historial de ejecutorias deficientes en estos exámenes que se queden en su casa durante los días de las pruebas. Tales fraudes escandalosos ya han surgido en California, New York, South Carolina, Virginia, y demás estados.

El poner la educación bajo el control de corporaciones distantes y el suprimirle autoridad a padres locales, maestros, administradores, juntas escolares, estudiantes y demás miembros de la comunidad también le hace daño a esta necesidad de un instrumento de evaluación confiable. Las alternativas a los exámenes estandarizados descritas en la sección V constituyen maneras mejores de evaluar logros académicos que también ofrecen un mayor grado de confiabilidad.

¿Acaso hay algo de bueno acerca de los exámenes estandarizados?

No mucho. A veces un estudiante de familia pobre o minoritaria que sale bien en un examen ha de recibir atención especial — y esto a pesar de que sus compañeros no recibirán igual trato. O una escuela con bajas puntuaciones puede que se vuelva el objeto de ayudas adicionales.

Las puntuaciones de estos exámenes han sido usadas para llamar la atención a la falta de buena enseñanza al alcance de niños con escasos recursos o de grupos minoritarios. En New York, estos exámenes se han usado para asegurarse que un mayor número de estudiantes sean atentidos adecuadamente en programas bilingües.

Los artículos examinadores de selección múltiple o de respuestas breves podrían ser usados como una pequeña parte del programa estatal de exámenes. Si tal programa tomase muestras de individuos y no fuera administrado en cada grado, podría proveer información útil sobre la confiabilidad de un examen sin hacerle daño o lastimar el plan de estudio.

En resumidas cuentas, el uso de exámenes estandarizados ha hecho muy poco para mejorar la educación de estudiantes individuales o de la sociedad. Sus muy pocos usos positivos han sido excedidos por el daño que infligen a una mayoría de los niños y de las escuelas.

BEST COPY AVAILABLE

V. MEJORES MANERAS DE EVALUAR ESTUDIANTES

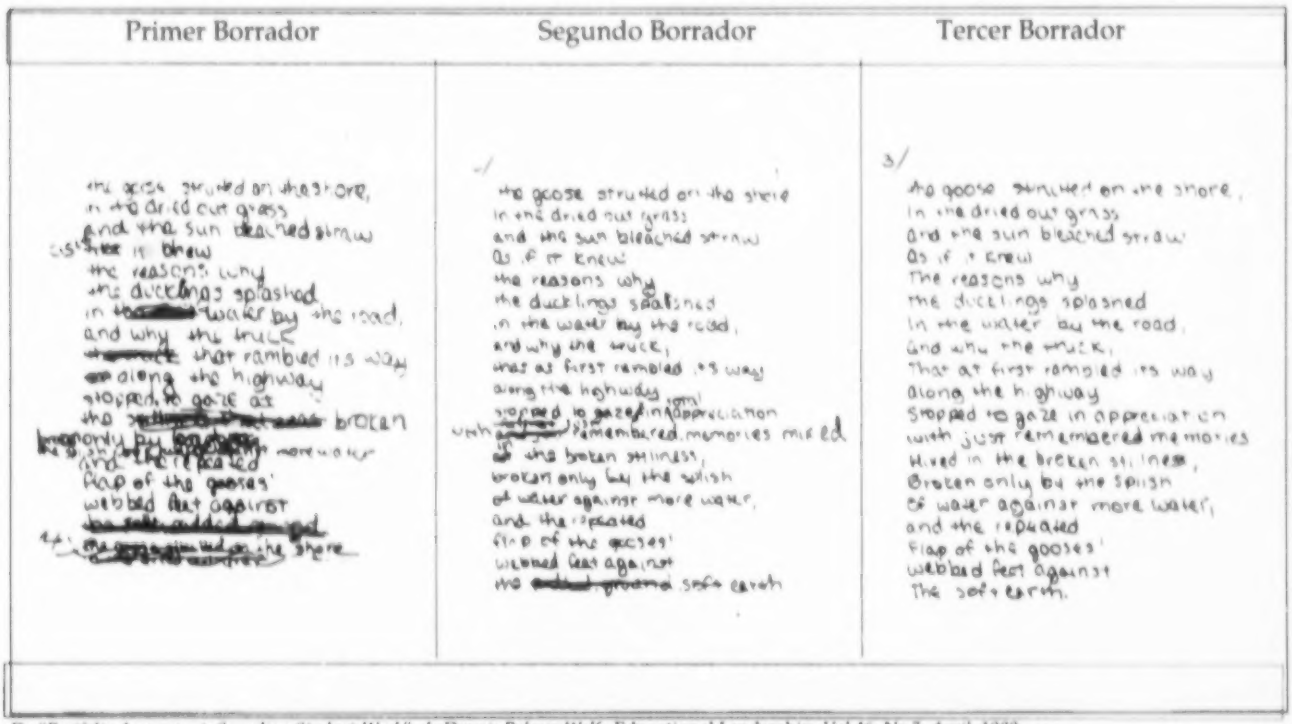
El propósito más importante de una evaluación debe ser ayudar al aprendizaje. Así pues, La Asociación de Juntas de Escuelas Públicas del Estado de Nueva York ha dicho: "El propósito principal de una política de evaluaciones debe ser la mejora en el aprendizaje del estudiante individual". El segundo propósito de la evaluación es proveer información a gente que están fuera de la escuela — padres, juntas escolares, legisladores — sobre la calidad del aprendizaje estudiantil.

Las pruebas estandarizadas de selección múltiple fracasan en ambos propósitos. Afortunadamente, ahora existen mejores maneras de evaluar a los niños. Estos métodos se conocen por nombres tales como **documentación, evaluación en base a tareas, evaluación auténtica o apropriada,** y **medición directa.**

Todas estas alternativas a las pruebas estandarizadas están basadas en una idea sencilla: debemos evaluar a los estudiantes justo cuando están ejecutando sus tareas y demostrando conocimiento del material que queremos que aprendan. Por ejemplo, si queremos que los estudiantes escriban con claridad, las escuelas deben examinar y corregir lo que escriben.

La **Observación y Documentación** son los métodos de evaluación más básicos. El maestro observa a la actividad de cada estudiante, toma notas, y combina esta información con muestras del trabajo del estudiante. Un ejemplo de la documentación es el seguir con cuidado como escribe un niño. En cuanto a la lectura, podemos anotar qué libros ha leido un estudiante, si él o ella los entiende lo suficientemente bien como para explicarlos con sus propias palabras, y si el estudiante puede hacer uso de la experiencia o conocimientos que contienen los libros.

También se pueden guardar muestras del trabajo de un alumno en un **portafolio.** Con respecto al escribir, este expediente puede contener



Primer Borrador	Segundo Borrador	Tercer Borrador
the geese strutted on the shore, in the dried out grass and the sun bleached straw as if it knew the reasons why the ducklings splashed in the water by the road, and why the truck that rambled its way along the highway stopped to gaze at the broken only by the and the repeated flap of the geese' webbed feet against the soft earth	2/ the geese strutted on the shore in the dried out grass and the sun bleached straw as if it knew the reasons why the ducklings splashed in the water by the road, and why the truck, that at first rambled its way along the highway stopped to gaze in appreciation with just remembered memories filled of the broken stillness, broken only by the splash of water against more water, and the repeated flap of the geese' webbed feet against the soft earth	3/ the goose strutted on the shore, in the dried out grass and the sun bleached straw as if it knew The reasons why the ducklings splashed in the water by the road, and why the truck, That at first rambled its way along the highway stopped to gaze in appreciation with just remembered memories filled in the broken stillness, Broken only by the splash of water against more water, and the repeated Flap of the geese's webbed feet against The soft earth.

De "Portfolio Assessment: Sampling Student Work", de Dennie Palmer Wolf. Educational Leadership, *Vol 46, No.7, April, 1989.*
Reimpreso con el permiso de la Association for Supervision and Curriculum Development. Copyright 1989 del ASCD. Se reservan todos los derechos.

BEST COPY AVAILABLE

varios borradores de un trabajo, con comentarios tanto del maestro como del estudiante. Los examinadores externos podrían mirar a una muestra de portafolios para comprobar si una clase entera está aprendiendo a un nivel satisfactorio. El estado de Vermont esta comenzando a usar evaluaciones de portafolios.

La documentación y los portafolios pueden ser utilizados para ayudar a los maestros a convertirse en mejores instructores. Este proceso también puede ser usado para protejer a los estudiantes de prejuicios de parte de sus maestros. Por ejemplo, otros maestros, administradores, y personas entrenadas de la comunidad pueden examinar directamente las muestras de trabajo estudiantil, observar los niveles de logros en diversas materias, y comparar la evaluación del maestro con el trabajo que ellos pueden ver en los portafolios. Si los observadores detectan problemas, podrían entrevistar al maestro y a los estudiantes, así como visitar algunas clases. El proceso evaluativo puede ser la base para trabajar en conjunto con los maestros con tal de mejorar a su enseñanza y sobreponer a cualquier prejuicio. Los maestros pueden utilizar parte de la información contenida en los portafolios para formar **listas de repaso** e **inventarios**. Estos métodos ayudan a almacenar de modo sistemático información sobre el progreso del estudiante.

Los **Exámenes de Respuesta Libre** son otra mejora. A los estudiantes se les puede pedir que, sin partir de indicaciones dadas por los examinadores, resuelvan por su cuenta a un problema matemático que contenga diversos pasos. En vez de tener que "adivinar" la respuesta entre una serie de contestaciones ofrecidas, tendrían que entender cuál es en realidad el problema, solucionarlo, y explicar por qué lo resolvieron de cierta manera y no otra. Por ejemplo, a una clase se le puede pedir que escriba un ensayo que intente persuadir a la junta escolar para que detenga a los exámenes innecesarios.

Estos exámenes pueden ayudar a los maestros al mostrarles cómo piensa el estudiante y cómo utiliza su conocimiento en diversas materias. Pueden ser un método para comprobar los resultados observados en el trabjo de portafolio. También pueden serles útiles a individuos fuera de la escuela. Arizona, California, Connecticut y Kentucky son algunos de los estados que van desarrollando exámenes de respuesta libre.

Ilustrado por Steve McCracken. Reimpreso con pemiso.

BEST COPY AVAILABLE

Se conoce como **ejecución** a un tipo de evaluación que por un lado es un examen, y por otro una documentación o portafolio. El llevar a cabo una obra de teatro, dar un recital musical, o competir en un evento atlético son todos ejemplos de ejecuciones. También lo es elaborar un **producto** como una mesa para un curso de carpintería o un dibujo para la clase de arte.

Las ejecuciones y los productos también se pueden usar para materias académicas. Así pues, los estudiantes pueden hacer investigaciones históricas, escribir sus hallazgos (un producto) y explicar y contestar preguntas sobre los mismos (una ejecución). El resultado final puede ser puntuado, quizás por un equipo de educadores, padres, y otros miembros de la comunidad. Cada estudiante puede recibir comentarios sobre su trabajo tanto durante como al final de su labor. Esto ya sucede con los proyectos exibidos en las ferias de ciencia estudiantil. Muchas escuelas incluso requieren que el estudiante cree a una ejecución, producto, o **exibición** de envergadura previo a la graduación. Todos estos métodos son formas tanto de examinar como de documentar. Pero contrariamente al tener que tomar exámenes de selección múltiple, estas pruebas pueden ser experiencias valiosas para los estudiantes.

La Calificación de Exámenes en Base a la Ejecución

No es difícil diseñar sistemas de puntuaciones que reemplazen a los exámenes estandarizados. Los jueces de los Juegos Olímpicos en eventos tales como el patinaje o la gimnasia le dan a los atletas una puntuación basada en la calidad y dificultad de su ejecución. La misma cosa puede ocurrir en los salones de clase. A los maestros se les puede entrenar para que den clasificaciones basadas en una escala común. Una puntuación de "6" puede ser dada a un trabajo escrito excelente, una de "5" a uno muy bueno, y así sucesivamente hasta un "1" para el trabajo muy mal escrito. De ésta forma se puede comparar a los estudiantes entre sí y ante un estandarte de excelencia. Hemos aprendido tanto de otros países que usan mucho a ensayos para evaluar a sus estudiantes, como de competencias atléticas, e igualmente de ferias científicas, que las puntuaciones que dan los calificadores entrenados generalmente concuerdan.

Se pueden usar sistemas parecidos para otras materias. Las evaluaciones del estudio de las ciencias pueden basarse en la puntuación de proyectos de investigación, experimentos de laboratorio, ensayos, y preguntas de respuesta libre. A través del proceso de realizar estas tareas, los estudiantes aprenden los hechos importantes y las reglas que son lo que intentan medir los exámenes de ahora. Aún más importante, ellos aprenden cómo pensar científicamente al tratar con problemas de tipo real.

Estas alternativas examinadoras no son perfectas, a pesar de que todas le enseñan al estudiante a pensar de forma más compleja y significativa. Los maestros pueden dedicarle demasiado tiempo al entreno de sus estudiantes para las tareas de los exámenes. O también pueden dar puntuaciones basadas en sus prejuicios y no en la ejecución de los estudiantes. Pero los padres concienzudos y evaluadores externos deben poder prevenir estos peligros. Es por eso que es muy importante incluir un sistema de cotejaciones que garantize la precisión e imparcialidad del examen.

Si son debidamente diseñados y supervisados, estos nuevos sistemas examinadores y evaluadores pueden mejorar la enseñanza y el aprendizaje. Los salones de clase podrán dejar de convertirse en campos de entreno para poder pasar exámenes. Habrá menos estudiantes alienados por materiales aburridos que son fomentados por los exámenes. Padres, educadores, y el público podrán ver mejor cuán efectivas son nuestras escuelas.

VI. LO QUE USTED PUEDE HACER

Usted puede ayudar a reducir el número de los exámenes y conseguir una evaluación auténtica para nuestras escuelas formando grupos, coaliciones, y campañas que confronten estos asuntos a nivel local o estatal.

Metas de la Campaña

Mientras más aprenda usted acerca de los exámenes y sus alternativas, más podrá ayudar. Temas de los cuales debe enterarse incluyen:

- ¿Cuáles son los exámenes dados en su distrito local y en su estado?
- ¿En qué grados se examina a los estudiantes?
- ¿Cómo son utilizados los resultados de los exámenes?

Al ser sumados a la información contenida en esta guía, estos conocimientos le ayudarán abogar más efectivamente. No permita que la falta de conocimiento o experiencia le impida comenzar a ser activo. Usted puede aprender lo que necesita saber.

Forme un Grupo o Coalición

Si usted desea efectuar cambios, debe trabajar con otras personas. Uniéndose a indviduos y grupos que comparten su opinión, le será más fácil correr la voz acerca de su causa y lograr sus metas.

Si usted enseña, puede comenzar hablando con otros maestros y maestras en su escuela acerca de cómo afectan los exámenes estandarizados a la enseñanza y los estudiantes. Usted puede llevar el asunto a su grupo o sindicato de maestros.

Como padre o madre, usted puede hablarle a amistades que sean padres de niños de edad escolar, o puede presentar el asunto en su grupo de padres local.

Si usted es un administrador o administradora, usted puede examinar cómo su escuela o distrito utiliza exámenes. Los principales pueden trabajar para reducir la presión de los exámenes y ayudar a introducir alternativas. Los superintendentes a menudo pueden reducir ellos mismos el número de los exámenes, investigar alternativas, y ayudar a maestros a comenzar su implementación.

Maestros, padres y administradores deben intentar discutir estos asuntos entre sí. Si usted puede incluir diferentes tipos de personas en su grupo o coalición, aumentará la probabilidad de que tenga éxito.

No es esencial que usted comienze con un grupo grande. Un grupo pequeño puede lograr mucho.

TIGRE por Bud Blake

SI LOS MANUFACTURADORES DE AUTOS RECLAMAN A LOS CARROS CON ERRORES...
Y LOS FABRICANTES DE JUGUETES RECLAMAN A LOS JUGUETES...
¿PORQUE NO PODEMOS NOSOTROS RECLAMAR A LOS EXÁMENES CON ERRORES?

Reimpreso con el permiso especial de *King Features Syndicate, Inc.*

BEST COPY AVAILABLE

En Denver City, Tejas, dos padres de niños de escuela elemental fueron el motor detrás de *Kids, Education and You* (*KEY* – Los Niños, La Educación, y Usted). Sus esfuerzos por detener al sistema de retención — a menudo en base a los resultados de exámenes — de niños en el pre-jardín y jardín de infantes condujo a cambios locales y luego ayudó persuadir a la Junta de Educación de Tejas para que prohibiese este sistema. La labor de KEY tuvo peso cuando se consiguió que la Junta dictaminase que los cursos transicionales entre el jardín de infantes y el primer grado constituyen una forma de retener a niños. *The Committee for Effective School Evaluation* (CESE – El Comité para la Evaluación Escolar Efectiva), un grupo de Arizona que persuadió al estado para que abandone a los requisitos de exámenes en los grados primero y duodécimo, fue guiado por tan sólo dos personas — un maestro y un padre.

Individuos o grupos que pueden apoyar sus esfuerzos incluyen: padres, maestros, administradores escolares, miembros de juntas escolares, negocios locales, legisladores locales o estatales, y organizaciones educacionales, de derechos civiles, y de comunidades. Planifique cómo alcanzar y persuadir a estos tipos de personas. Averigüe quién tiene mayor influencia con aquellos que toman decisiones en sitios tales como las juntas escolares. Es sobre todo a estos individuos que usted debe persuadir.

Establezca Metas Claras

Al formar un grupo o coalición, usted debe ponerse de acuerdo con una meta o propósito general. En Nueva York, por ejemplo, la gente que formó la *New York Campaign for Genuine Accountability in Education* (Campaña de Nueva York Pro-Genuina Responsabilidad en la Educación) escribió una Declaración para Urgir la Prohibición de los Exámenes Estandarizados para Niños Menores. El ponerse de acuerdo con una meta desde el principio le brinda un enfoque a su plan de acción mientras va aumentando su membresía y el número de sus actividades. (Véanse a las páginas 27-28 para la Declaración Nacional Pro-Genuina Responsabilidad en la Educación, firmada por más de 75 organizaciones.)

Usted puede desarrollar metas de plazo corto, medio, o largo. Una meta de plazo corto podría ser el reunirse con la junta escolar de su distrito para presentar su posición y averiguar el punto de vista de la junta. Una meta de plazo medio podría ser reducir o eliminar a los exámenes estandarizados en los primeros grados. Una meta de plazo largo podría ser cambiar las leyes estatales con respecto a los exámenes.

Cada estado, distrito, o escuela es diferente. Por tanto, las metas deben desarrollarse en base a las necesidades específicas de su comunidad o estado. La Campaña de Nueva York, por ejemplo, se está enfocando en el examen de niños menores ya que la mayoría de sus distritos examina en los primeros grados, a pesar de que el estado no exige los exámenes hasta el tercer grado.

Haga uso de las experiencias y los recursos que ya están disponibles. Las personas envueltas en la Campaña de Nueva York aprendieron de la Campaña de Carolina del Norte la cual logró detener a los exámenes en el primer y segundo grado. También obtuvieron conocimientos de parte de las experiencias de aquellos que habían trabajado a través de los años en Nueva York para cambiar las formas con las cuales se usan los exámenes de admisión para las universidades.

Al escoger sus metas, y en especial las de corto plazo, seleccione a aquellas que usted tenga posibilidad de poder realizar: "No hay nada tan exitoso como el éxito." Sin embargo, asegúrese de que sus metas son lo suficientemente importantes como para recompensar a sus esfuerzos.

Desarrolle una Estrategia—Invite Participación

Una estrategia es un plan para alcanzar su meta. En ella se incluyen todos los pasos importantes que usted debe tomar para ganar. Por ejemplo, el

Atlantic Center for Research in Education (ACRE — Centro Atlántico para la Investigación en la Educación) tuvo como meta el abolir los exámenes dictaminados a nivel estatal para los grados primero y segundo. Esto requirió el paso de una legislación. Para ganar, ACRE tuvo que documentar problemas, atender reuniones, ayudar organizar a maestros de niños menores, y persuadir a algunos — y luego a muchos — legisladores para que diesen su apoyo.

Para lograr éxito siempre es necesario educar a individuos, grupos y legisladores claves, así como al público en general. Esto puede lograrse de varias maneras: con folletos, presentaciones públicas, hojas informativas, cartas a editores, reuniones, hojas sueltas, videos, comparecencias en la radio y la televisión tales como en programas de discusión — todo esto provee formas de difundirle información al público.

Es importante saber qué materiales o argumentos usar con qué gente. ACRE encontró que los argumentos más persuasivos con los legisladores fueron: 1) los niños menores son malos tomadores de exámenes, de modo que los resultados de los exámenes no son precisos 2) los exámenes estandarizados ignoran a metas educativas muy importantes para niños menores, que no pueden ser medidas con pruebas de papel y lápiz. En su área los mejores argumentos pueden ser distintos.

En última instancia, los activistas de la reforma de exámenes necesitan convencer a los oficiales públicos. Al informar y presionar a estos encargados a nivel local y estatal para que tomen decisiones, usted puede cambiar la política pública. Aprenda como se toman las decisiones. Encuentre grupos que le apoyen y que conozcan el proceso local y estatal de toma de decisiones.

Sea Flexible

La preparación para problemas que puedan surgir debe formar parte de la selección de metas y planes. Sin embargo, a menudo ocurre algo inesperado o se hacen disponibles nuevas informaciones que pueden afectar sus planes originales.

El compromiso puede ser apropiado. Por ejemplo, el grupo de Arizona acceptó un plan para que el estado detuviese los exámenes en el grado primero y duodécimo a pesar de que el plan original era detener los exámenes también en el grado segundo. Tuvieron que elegir entre este cambio en la legislación, o ninguno.

Si su estrategia no está funcionando, se necesita un plan completamente nuevo. Puede que tenga que hacer ajustes mientras va trabajando. De todos modos, de hacerse necesaria, la flexibilidad facilita a los ajustes y le permite hacer cambios en sus planes.

Prepárese para un Camino Largo

Sistemas que han existido por mucho tiempo, tales como la burocracia educacional, no pueden cambiarse fácilmente o de un día para otro. Al darse cuenta de que los resultados finales de su trabajo pueden tomar varios años, evitará usted las trampas que pueden surgir de expectaciones poco reales.

Si desea información, materiales (tales como hojas informativas), o asistencia para comenzar, póngase en contacto con la oficina de FairTest en la 342 Broadway, Cambridge, MA 02139; (617) 864-4810.

APÉNDICE: Recursos que Usted Puede Utilizar

Bibliografía

Los siguientes recursos están publicados en inglés, y pueden ser obtenidos a través de las agencias aquí citadas:

FairTest, 342 Broadway, Cambridge, MA 02139.

- *Fallout from the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools* (3ra edición, 1990). Contiene información y referencias que apoyan a esta guía. $8.95.
- *FairTest Examiner*. Boletín trimestral sobre la reforma de exámenes. $15 al año para individuos; $25 al año para instituciones.
- *Standing Up to the SAT*. (1989) Contiene información sobre la reforma de exámenes y también sobre cómo prepararse para saber tomarlos, todo en un mismo libro. $6.95.

FairTest también dispone de una bibliografía anotada sobre los exámenes en las escuelas privadas (incluida en el *Fallout*) que cubre todos los diversos problemas de las pruebas. Y FairTest provee *Fact Sheets* (hojas informativas) sobre los Exámenes K-12 y los de Admisión para ir a la Universidad. Use el formulario en la página 29.

NYPIRC, 9 Murray St., New York, NY 10007.

- *Failing Our Children: How Standardized Tests Damage New York's Youngest Students.*(1989) $5.00. Use el formulario en la página 29.

Otras lecturas

- Ascher, C. "Assessing Bilingual Students for Placement and Instruction," ERIC Clearinghouse on Urban Education *Digest*. (ERIC/CUE, Box 40, Teachers College, Columbia U., NY, NY 10027.) Gratis. Resumen, con referencias, de problemas en el examen de estudiantes bilingües.
- *Network (Back to School 1990): High Scores vs. High Achievement: Concerns About Standardized Testing in America's School Improvement Movement"*. (National Committee for Citizens in Education, 1080 Little Patuxent Parkway, Suite 301, Columbia, MD 21044.)
- *Rethinking Schools* (Jan/Feb 1989) Léase la sección "Focus on Testing". (P.O. Box 93371, Milwaukee, WI 53202; $2.00.)
- *National Association for the Education of Young Children* (NAEYC – Asociación Nacional para la Educación de Niños Menores) *Testing of Young Children: Concerns and Cautions* (1988: NAEYC, 1834 Connecticut Avenue, N.W. Washington, DC 20009.) $.50 cada uno o 100 por $10.
- Kamii, C., ed. *Achievement Testing in the Early Grades: The Games Grown-ups Play* (1990: véase la dirección de la NAEYC, citada arriba) $8.00.
- *Educational Leadership, April 1989: "Redirecting Assessment"*. 17 artículos sobre la evaluación auténtica. (Order Processing, ASCD, 1250 N. Pitt St., Alexandria, VA 22314.) $5.
- *Phi Delta Kappan, May 1989: Special Section on Testing* – contiene un artículo de FairTest y otros sobre la evaluación auténtica. (P.O.Box 789, Bloomington, IN, 47402.)

Los Exámenes Estandarizados y Nuestros Niños

Copias adicionales de esta *Guía para la Reforma de Exámenes* se pueden obtener por correo de Comuníquese con FairTest si quiere obtener derechos para reimprimir grandes cantidades. **Esta guía también esta disponible en inglés, al mismo precio, a los siguientes precios:**

Copia individual: $3.00
5 copias: $10.00
10 copias: $15.00
50 copias: $50.00
100 copias: $75.00

Use el formulario en la página 29.

APÉNDICE: Significado Real de Términos Examinadores

Las pruebas están repletas de términos que no les son familiares a muchas personas. Algunas palabras no tienen el mismo significado al ser utilizadas en los exámenes que el que tienen en su usanza común. Con tal de que usted pueda entender mejor a estas palabras al hablar con maestros, administradores, juntas escolares, y otros, he aquí un diccionario de términos básicos usados en los exámenes. Provee equivalencias con el inglés y con algunos términos del español:

Achievement Tests / **Exámenes de Logros**: pretenden medir cuánto ha aprendido el estudiante en clases tales como de lectura, deletreo, ciencia, o matemáticas. También son conocidos como **"de conocimientos (o de logros) por materia"**.

Aptitude Tests / **Exámenes de Aptitud**: tratan de predecir la habilidad de un estudiante para hacer ciertos tipos de trabajo escolar; a veces se les llama exámenes **"de habilidades cognoscitivas"**; las pruebas de I.Q. pueden ser usadas con este propósito.

Competency Tests / **Exámenes de Competencia**: están diseñados para determinar si un estudiante ha cumplido con el estandarte mínimo de destrezas y conocimiento en una área de estudio específica.

Criterion-Referenced Tests / **Exámenes Referidos a un Criterio**: miden cuán bien cumple un estudiante con los requisitos establecidos por el examinador. También se les conoce como **"evaluaciones de un proceso experimental."**

Cut Score / **Puntuación Mínima**: es el nivel mínimo que debe lograr el que toma una prueba para lograr "pasarla" (aprobarla).

Developmental Screening Tests / **Exámenes de Discernimiento de Desarrollo**: son dados para identificar a los estudiantes que pueden tener problemas que necesiten la ayuda especial de un especialista, como por ejemplo un doctor.

High Stakes Tests / **Exámenes de Alta Transcendencia**: son pruebas usadas para hacer decisiones educacionales importantes y que son vistas como importantes por los estudiantes y educadores.

IQ (Intelligence Quotient) Tests / **Exámenes de I.Q. (C.I., Cociente Intelectual)**: supuestamente miden la inteligencia natural de una persona al asumir que un sólo conjunto de habilidades heredadas puede ser medido por un examen.

Measurement Error / **Error de Medición**: son un medio por el cual se indica cuán confiable o no es un examen; todos los exámenes tienen un error de medición.

Norm-Referenced Tests / **Exámenes de Referencia Normativa** También se les conoce como **"calibrados por la población"**: son diseñados para que las puntuaciones de los estudiantes que los tomen recaigan a lo largo de una curva "normal" o con "forma de campana".

Reliability / **Confiabilidad**: calcula cuánto se aproximaría un exámen a producir los mismos resultados si fuese administrado a un mismo estudiante, bajo las mismas condiciones, pero en repetidas ocasiones.

Sampling / **Toma de Muestras**: es una forma de obtener información acerca de un grupo examinando a tan sólo unos cuantos miembros del grupo o dándole a todos los estudiantes sólo partes pequeñas del exámen.

School Readiness Tests / **Exámenes de Predisposición Escolar**: son utilizados para determinar si un niño esta "dispuesto" a comenzar el jardín de la infancia o el primer grado. También se conocen como **"pruebas de madurez para el aprendizaje"**

True Score / **Puntuación Verdadera**: la puntuación que debería obtener el estudiante si el examen fuera completamente confiable.

Validity / **Validez**: define cuán precisamente ha medido un exámen a lo que sus promotores sostienen que mide; la *validez de contenido* mide cuán bien cubre el examen al contenido de su materia; la *validez predictiva* mide cuán bien prognostica el examen al desempeño escolar de un estudiante.

Declaración de la Campaña Pro-Genuina Responsabilidad en la Educación

Nosotros, los miembros de la Campaña Pro-Genuina Responsabilidad en la Educación, creemos que el gran énfasis en los exámenes estandarizados y de selección múltiple como medida principal de habilidad o logros educacionales obstruye a las reformas educativas necesarias en nuestra nación. Hemos preparado esta Declaración con tal de urgir cambios fundamentales en la evaluación.

Es de importancia crítica que los oficiales políticos y educacionales, de todos los niveles, reconozcan que los exámenes estandarizados y de selección múltiple no son un medio adecuado para medir al progreso educacional, y que tampoco es una meta ápropiada el esfuerzo por elevar los resultados de los mismos. Existen otras medidas evaluadoras que son educacionalmente sanas y que ofrecen una base para mejorar la enseñanza, informar al público, y medir el progreso en el aprendizaje, proveyendo así una responsabilidad genuina.

Los estados, distritos locales, y el gobierno federal deben comprometerse a la meta de irse apartando de una dependencia en los exámenes estandarizados y de selección múltiple, e ir moviéndose hacia métodos evaluadores del progreso y logros educacionales que ofrezcan alternativas y sean auténticos y apropiados. Al hacer esto:

- Los líderes educacionales y políticos *deben* reclamar el uso de una evaluación auténtica y apropiada en los salones de clase y evaluar al conjunto del progreso y logros estudiantiles; *deben* pedirle a sus estados, distritos locales, y programas con fondos federales, tales como *Chapter I*, que implementen las alternativas evaluadoras que ya están disponibles, y que continuen desarrollando nuevos métodos para evaluar a estudiantes y programas.
- Estos líderes *no deben* reclamara un incremento en las puntuaciones de los exámenes estandarizados como meta de una reforma educacional; *no deben* reclamar nuevos o adicionales métodos examinadores por medio de pruebas estandarizadas y de selección múltiple; *no deben* depender de los exámenes estandarizados existentes como indicadores del progreso educacional; y *no deben* depender de los exámenes estandarizados para hacer decisiones en torno a la colocación de estudiantes dentro de programas. Estos exámenes dirigen mal al plan de estudios, son medidas educativas inadequadas, no son útiles en la instrucción, a menudo son demasiado imprecisos para servir como base para decisiones sobre individuos — y pueden ser reemplazados por métodos evaluadores auténticos.

Las alternativas que nosotros apoyamos se conocen por diversos nombres: auténticas, apropiadas, en base a ejecuciones, y directas. Son evaluaciones que se basan en preguntarle a los estudiantes que demuestren lo que nosotros queremos que ellos sepan y puedan hacer. Por ejemplo, si los estudiantes deben escribir bien, debemos examinar cómo escriben. La evaluación apropiada motiva al pensamiento real e incluye a observaciones por parte de maestros, ejecuciones, exibiciones, portafolios del trabajo estudiantil, listas de prueba, y preguntas de respuesta libre. Estas evaluaciones proveen información que es útil en la instrucción, pueden evitar los prejuicios raciales y culturales de muchos de los exámenes estandarizados, y pueden ponerse en formas que provean información útil y una genuina responsabilidad ante el gobierno local, estatal, y nacional.

Los líderes políticos y educacionales de todos los niveles deben proponerse un plazo fijo para ir eliminando a la mayoría de los usos de los exámenes estandarizados y de selección múltiple, reemplazándolos con alternativas confiables, válidas, y educacionalmente apropiadas. En el periódo previo a la implementación de estas alternativas, el uso y dependencia en los exámenes estandarizados y de selección múltiple debe ser reducido con la mayor intensidad y rapidez posibles.

Lista de Firmadores

"Declaración de la Campaña para la Responsabilidad Genuina en la Educación"

Advocates for Children of New York
American Association of Colleges for Teacher Education
American Federation of Teachers
American Reading Council
Angelo N. Ancheta, Asian Law Alliance*
APPLE Corps., Inc.
Association for Supervision and Curriculum Development
Bank Street College of Education, New York
Rims Barber, Mississippi Human Services Agenda
Leonard Beckum, Dean, College of Education, CUNY
Association for Women in Science
Diana Caballero, Puerto Rican/Latino Education Roundtable, NYC*
Center for Law and Social Policy
Center for Women Policy Studies
Council for Basic Education
Robert W. Covert, Past President, American Evaluation Association*
Harold E. Dent, Vice-President, Psychological and Human Resources Consultants, Inc.*
Educational Law Center, Inc.
Pablo Eisenberg, President, Center for Community Change*
Federation of Organizations for Professional Women
Foxfire Fund, Inc.
Fund for the Feminist Majority
Howard Gardner, Project Zero, Harvard University*
Genesee Valley Developmental Learning Group, New York
Leslie A. Hart, Brain-Compatible Education Associates
La Donna Harris, Americans for Indian Opportunity*
Asa Hilliard, Professor of Education, Georgia State University*
Holt Associates/*Growing Without Schooling*
Institute for Learning and Teaching
International Reading Association
Kentucky Association on Children Under Six
KEY-Kids, Education and You
A. Gay Kingman, National Congress of American Indians*
Louisiana Association on Children Under Six
Massachusetts Advocacy Center
Matsushita Foundation, Inc.
Deborah Meier, Principal, Central Park East Secondary School, NYC*
Samuel J. Meisels, Professor of Education, U of Michigan*
Sara E. Melendez, Vice-Provost, University of Bridgeport*
Mexican American Women's National Association
National Association for the Advancement of Colored People
National Center for Fair & Open Testing (FairTest)
National Coalition of Advocates for Students
National Coalition of Title I/Chapter I Parents
National Council of Teachers of English
National Council for Social Studies
National Education Association
National Indian Youth Council, Inc.
National Organization for Women--New York City Chapter
National Parent Teacher Association
National School Boards Association
National Women's Law Center
National Women's Political Caucus
Fred M. Newmann, National Center on Effective Secondary Schools--University of Wisconsin*
New York Public Interest Research Group
NOW Legal Defense and Education Fund
Organization of Chinese American Women
Vito Perrone, Harvard University--Graduate School of Education*
Project Equality, Inc., Kansas City, MO
Puerto Rican Legal Defense and Education Fund, Inc.
Rochester (NY) Teachers Association (AFT)
Lori Rubenstein, Partnership for Democracy*
William Saunders, National Alliance of Black School Educators*
William V. Schipper, Executive Director, Natl. Assn. of State Directors of Special Education*
Hilton Smith, Coordinator, Foxfire Teacher Outreach Programs*
Southern Association on Children Under Six
Southern Christian Leadership Conference
Southern Regional Council, Inc.
Adria Steinberg, Harvard Education Letter*
United Federation of Teachers (New York City)
United States Student Association
Dorothy J. Watson, President, Whole Language Umbrella*
Paul Weckstein, Center for Law and Education*
Grant Wiggins, Center on Learning, Assessment and School Structures (CLASS)
Arthur E. Wise, NCATE*
Women's Research and Education Institute

***Organización tan sólo para propósitos de identificación.**

Muchos otros individuos, en especial maestros y maestras, han endorsado a esta declaración.

Publicaciones FairTest

Standardised Tests and Our Children:
EXAMINER
Fallout From the Testing Explosion
NONE OF THE ABOVE
Behind the Myth of Scholastic Aptitude
Sex Bias in College Admissions Tests: Why Women Lose Out
STANDING UP TO THE SAT
FAIRTEST

Todas las publicaciones siguientes son en inglés, salvo las dos primeras guías, que también están disponibles en español:

Standardized Tests and Our Children: A Guide to Testing Reform, de FairTest (33pp.)..1-$3.00, 5-$10.00; 10-$15.00; 50/$50.00; 100/75.00. Las versiones en español e inglés pueden ser combinadas para los precios en conjunto.

Inglés _____
Español _____

Standardized Tests and Our Children: A Guide to Testing Reform in New York, de FairTest & NYPIRC (30pp.)...los mismos precios que los de arriba. Precios en conjunto disponibles de: NYPIRC-184 Washington Ave.; Albany, NY 12210

Inglés _____
Español _____

FairTest Examiner, publicado quatrimestralmente...$15.00/año (individuos), $25.00/año (instituciones) _____

Publicaciones previas del =FairTest Examiner, $4.00 cada: Vol. __ No. __ (Todos 13 por $40.00) _____

Standing Up to the SAT, por John Weiss, Barbara Beckwith y Bob Schaeffer (192pp.)...$6.95 _____

Fallout From the Testing Explosion: How 100 Million Standardized Exams Undermine Equity and Excellence in America's Public Schools, de Monty Neill y Noe Medina (45pp.)...$8.95 _____

Sex Bias in College Admissions Tests: Why Women Lose Out, de Phyllis Rosser y encargados de FairTest, 3ra edición (60pp.)...$7.95 _____

The Reign of ETS, de Allan Nairn y Ralph Nader (550pp.)...$30.00 _____

Beyond Standardized Tests: Admissions Alternatives That Work, de Amy Allina con encargados de FairTest (18pp.)...$5.50 _____

None of the Above, Behind The Myth of Scholastic Aptitude Tests, de David Owen (327pp.)...$5.95 _____

Contribuciones a FairTest

¡Estoy de acuerdo en que no hay mejor manera de formar parte del movimiento para la reforma de exámenes que apoyar a FairTest!

Asociado de FairTest ($25.00—incluye una suscripción gratis al *FairTest Examiner*_____

Donante a FairTest ($50.00—incluye la suscripción al *Examiner* y la publicación que usted elija)_____

Mantenedor de FairTest ($100.00 o más—incluye lo de arriba más un descuento de 20% en las publicaciones)_____

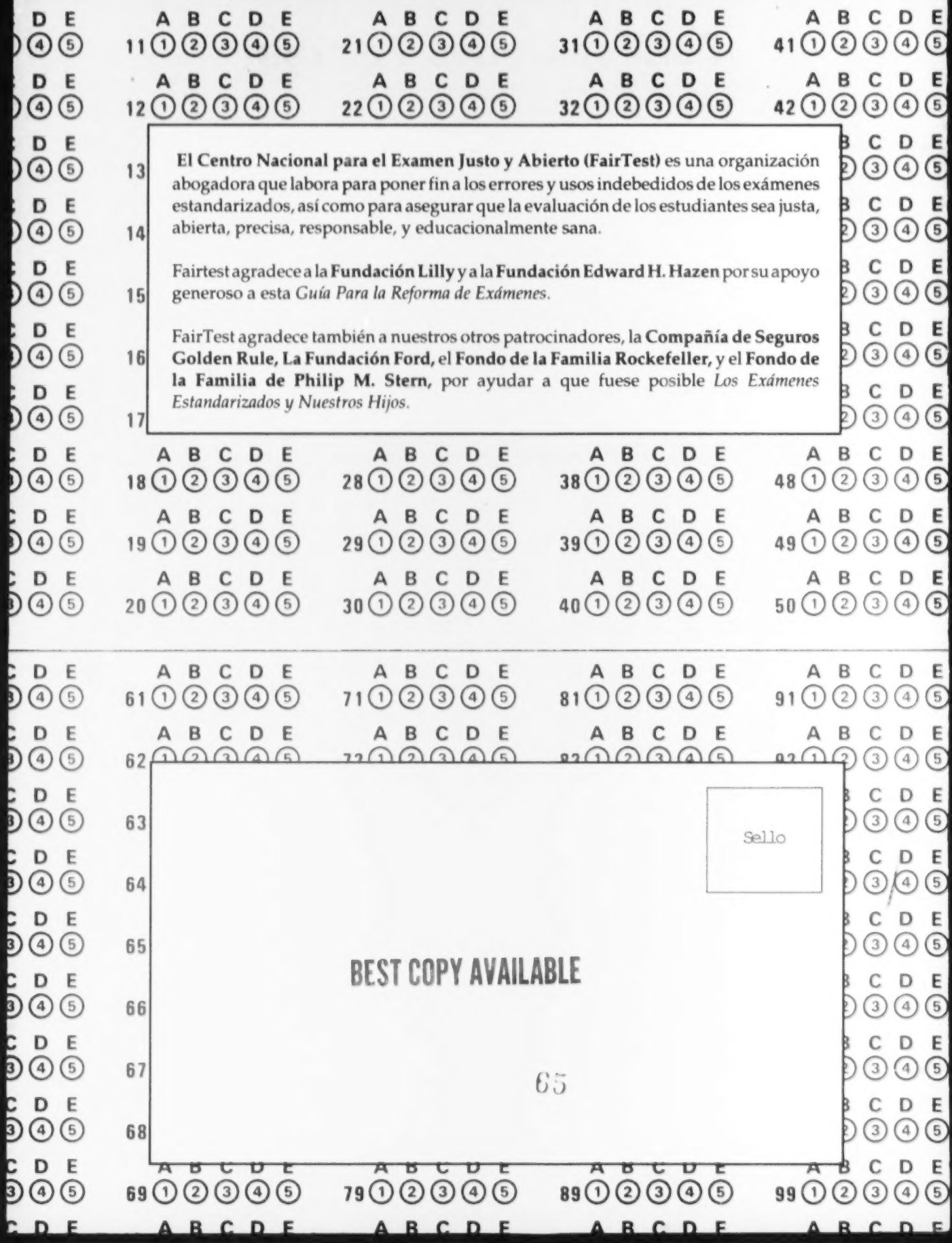
Otra_____

Publicaciones en Total _____
Contribuciones en Total _____
Total del Cheque _____

Favor de hacer sus cheques pagables a FairTest y de enviarlos a: FairTest, 342 Broadway, Cambridge, MA O2139

Nombre _____
Dirección _____
Ciudad _____
Estado _____ Código Postal _____

**Las contribuciones a Fairtest son netamente deducibles de impuestos, excepto por el valor de las publicaciones recibidas.*



El Centro Nacional para el Examen Justo y Abierto (FairTest) es una organización abogadora que labora para poner fin a los errores y usos indebedidos de los exámenes estandarizados, así como para asegurar que la evaluación de los estudiantes sea justa, abierta, precisa, responsable, y educacionalmente sana.

Fairtest agradece a la **Fundación Lilly** y a la **Fundación Edward H. Hazen** por su apoyo generoso a esta *Guía Para la Reforma de Exámenes*.

FairTest agradece también a nuestros otros patrocinadores, la **Compañía de Seguros Golden Rule, La Fundación Ford,** el **Fondo de la Familia Rockefeller,** y el **Fondo de la Familia de Philip M. Stern,** por ayudar a que fuese posible *Los Exámenes Estandarizados y Nuestros Hijos*.

Sello

BEST COPY AVAILABLE